सत्यमेव जयते

INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI

22078

I.A.R.I.6.
GIP NLK—H-3 I.A.R.I.—10-5-55—15,000

PHYSIOLOGICAL REVIEWS

VOLUME 27

BALTIMORE, MD.
1947

22078

CONTENTS

No. 1. January, 1947

No. 2. April, 1947

No. 3. July, 1947

No. 4. October, 1947

PHYSIOLOGICAL REVIEWS

VOL. 27 JANUARY, 1947 No.

THE REGULATION OF THE RESPIRATORY MOVEMENTS BY PERIPHERAL CHEMO-RECEPTORS

AUGUST PI-SUÑER

Instituto de Medicina Experimental—Universidad de Caracas

1. The regulation of the respiratory movements by chemical influences on the respiratory centers has been demonstrated repeatedly since the classical observations of Rosenthal (1880) on the respiratory effects of oxygen-lack; of Pflüger (1868) on the influence of metabolites accumulated in the blood; and of Hermann (1870) on the action of carbon dioxide and the variations in excitability of the centers which depend upon the oxygen tension of the blood. Of particular importance are the experiments of Fredericq (1890), using for the first time the method of crossed circulation, and the more recent work of Haldane and Priestley (1905), Haldane and Poulton (1909), Henderson (1908–38), Winterstein (1911), Hasselbalch (1912), Hooker, Wilson and Connett (1917), Gesell (1925), Dautrebande (1930), Winterstein and Frühling (1934), and many others. The increase in the concentration of carbon dioxide in the alveolar air, in the blood and in the cells themselves which form these centers stimulates the pulmonary ventilation.

This action of carbon dioxide, or of oxygen-lack, on the various types of respiration in air or in water is found in the most diverse species of animals. It represents a universal biological phenomenon. It is evident in mammals, birds, amphibia and reptiles. Westerlund (1906), Olthoff (1936) and Powers and Clark (1942) observed it in fishes; Fox and Johnson (1934), Peters (1938) and Lindroth (1938) in crustacea; Winterstein (1925) in certain cephalopods, etc. It is interesting to point out that in aquatic respiration a low oxygen tension is more effective than the concentration of carbon dioxide in increasing the respiration (Meyer, 1935; Heerdt and Krijgsman, 1939) an effect which is not surprising considering the physical conditions of the medium. These important questions have been examined by Krogh (1941).

The control of pulmonary or branchial ventilation occurs through the action of the centers which are affected by the gas content of the internal environment in relation to the composition of the external environment. But the possibility of peripheral chemo-receptors which may evoke reflexes adequate for the respiratory needs must also be considered.

"For various reasons," I wrote in 1918, "we consider that in *addition to the central chemical influences* there must be a peripheral regulatory factor acting through receptors which are also sensitive to chemical changes. Just as the vagus transmits afferent impulses in response to mechanical stimuli at the level of the lung which have reflex effects, so there must be pheripheral receptors which are excited by variations in the concentration of carbon dioxide."

This hypothesis has classical antecedents. Donders (1853), Berns (1870) and Traube (1871) thought that the increase of carbon dioxide in the air contained in the lungs must be concerned in the regulation of the respiratory movements.

Nevertheless, because of the lack of sufficient experimental proof, because of the study of the mechanical reflexes begun by Hering and Breuer (1868) and continued by Gad (1880), Zuntz and Geppert (1888), Head (1889) and many others, and because of the very important work of Haldane and his school and of those previously mentioned on the effect of the gas content of the blood on the functions of the respiratory centers which regulate the movements, no attention was paid to the peripheral chemical control of the respiration.

Starling stated in his *Principles of Human Physiology* (1915): "If . . . we succeed in altering the tensions of the two gases (oxygen and carbon dioxide) in the alveolar air we may assume that the tensions of the gases in the arterial blood leaving the lungs are altered in the same ratio." Through the action of chemoreceptors in the terminal portions of the respiratory apparatus, the composition of the alveolar air and therefore of the gas content of the blood is adjusted with the result that the ventilation would always be what the circumstances require. Only high concentrations of carbon dioxide in the inspired air, which increase the ventilation enormously, are able to raise the concentration of carbon dioxide in the alveolar air appreciably. The great speed of the responses, and the exactitude with which compensation is produced, lead one to suppose that there are reflexes evoked by chemical excitation from the respiratory apparatus which would operate previous to any humoral influence.

The experiments of Haldane (1922) show that if air with increasing quantities of carbon dioxide is breathed, it is possible for the ventilation rate to be doubled with practically no change in the composition of the alveolar air. Campbell, Douglas and Hobson (1914) observed that an increase of 2.5 mm. Hg in the carbon dioxide tension is sufficient to bring the ventilation to 10 liters per minute. These observations have been confirmed repeatedly (Campbell, Douglas, Haldane and Hobson, 1913; Douglas and Harvard, 1932; Barcroft and Margaria, 1931).

Nor, on the other hand, does rarefaction of the air change the gas content of the arterial blood appreciably, as has been known since the work of Fränkel and Geppert (1883): at a pressure of 410 mm. Hg there is no change in the gas content of the blood; nevertheless the respiratory dynamics have been modified.

The results obtained by Scott (1908) furnish very interesting data: without the vagi, the control of the respiration is uncertain, the reactions are slower and last longer after a carbon dioxide mixture has been breathed, and in general the adjustment of the respiration to physiological requirements is inexact. If the respiratory movements depended solely upon the humoral influence on the centers, and the vagus were limited to the rôle of a sensory nerve for mechanical receptors, there would not be such marked differences in the respiratory reactions to inspired carbon dioxide when the vagus is intact and when it has been cut. Our experiments designed to clarify this problem have been numerous and of various types.

At first we studied (1918) the behavior of the respiration in vagotomized dogs exposed to the same concentrations of carbon dioxide in air, before and after the operation. Observations were made immediately after the section of the vagi or some hours later. The differences with respect to the normal were always

noteworthy in that they showed lack of precision and delay in motor reactions when the vagal innervation was absent. Our conclusions were: 1. The constancy of the proportion of CO_2 in the alveolar air contradicts the hypothesis that the respiratory stimulation is exclusively of central origin evoked by the excess of CO_2 or lack of O_2 in the blood. 2. The lung and the bronchial ramifications are sensitive to different chemical stimuli, which can evoke reflexes. 3. In addition to the well known action on the respiratory centers, there is exerted a parallel or perhaps a previous peripheral influence due to the excitation of end-organs which are sensitive to stimuli of chemical nature by the CO_2 contained in the inspired air.

The following year in collaboration with Bellido (1919–21) and as a more evident demonstration we devised a crossed-circulation technique, otherwise known as the "dog with two heads." The central stumps of the carotids and jugulars of the donor dog "A" are joined to the cephalic stumps of the corresponding vessels of "B," the dog in which the experiment is performed. The vertebral arteries of the latter dog are also ligated, and the animals are chosen so that "A" is sufficiently larger than "B" and its carotid pressure higher, thus preventing the head of the recipient from receiving its own blood through the intravertebral plexuses. C. Heymans and Ladon in 1925 and J. F. and C. Heymans in 1926 proposed a definitive method similar to ours except that they cut off the head of "B" (method of the isolated head) and record respiratory movements of the larynx and floor of the mouth. With Puche (1930) we studied various ways of registering these movements with the purpose of improving the recording. Isolation of the head assures that only the blood of the donor dog reaches it, because it is connected to the trunk only by the vagi.

When the respiratory centers of "B" are perfused by either of the two procedures with the blood of "A," which is breathing normally or in some of the experiments is given artificial respiration, and "B" or its trunk which has been separated from the head is made to breathe air with CO_2, an increase in the frequency and depth of respiration is observed.

With Puche and Raventós (1930) we used also the technique of the isolated head. As in our first experiments of 1919 and succeeding years, if after decapitating "B," leaving the head which is connected with the trunk only by way of the vagi, we caused the trunk to breathe air with CO_2, the head responded with an increased intensity of the respiratory movements, despite the fact that it was being perfused with normal arterial blood from the donor dog "A."

The researches of J. F. and C. Heymans (1926–28) showed that the head, besides responding to mechanical stimulation of the pulmonary vagal endings (Hering-Breuer reflex), also responds to peripheral chemical stimuli. A reflex apnea is produced by over-ventilation as described by Hering and Breuer (1868) and afterwards studied by Baglioni (1903), Foa (1909–11), Githens and Meltzer (1914), Joseph (1922), Puche (1923), Meek (1923) and others. This apnea has been generally attributed to over-distention of the lungs, but Heymans and Heymans state that the major factor in producing it is the peripheral sensitivity to the concentration of CO_2 in the blood. Such a sensitivity was suggested by

Luciani (1888) and Bordoni (1888). Anemia of the trunk with asphyxia of the tissues causes acceleration of breathing, thus confirming the possible peripheral origin of the dyspnea in addition to its central origin, a thesis which had been proposed earlier by Francois Franck (1890), Hoffmann (1900), Porter and Newburgh (1916–17) and Dunn (1920). Heymans and Heymans were able to demonstrate also the relation between the cardio-aortic pressure and the breath rhythm, and, a fact which interests us particularly, respiratory responses to changes in the composition of the inspired air and in the gas content of the blood circulating through the heart and great vessels.

Heymans, Bouckaert and Regniers (1933) write: "J. F. and C. Heymans observed that intense hyperventilation of the trunk of dog 'B' causes reflex apnea of the isolated head of the same animal; on the other hand, progressive asphyxia of the trunk by withholding artificial respiration *or the administration of air with carbon dioxide,* causes the reappearance of respiratory movements of the head followed by a progressive increase in their amplitude as the asphyxia becomes more severe." Heymans and Heymans wrote in 1927: "These experiments taken together show that the vagi contribute to the reflex regulation of the activity of the respiratory center according to the peripheral respiratory and circulatory conditions. The respiratory center can be excited in the same way by a peripheral state of asphyxia or anoxemia as by a central state." Having demonstrated the existence of respiratory reflexes evoked by chemical stimuli Heymans and Heymans (1927) then set out to discover the peripheral origin of the respiratory tone and of the vagal influence on the respiration. For this purpose they performed various series of experiments demonstrating that the cardio-aortic intraceptive zone is the place of origin not only of circulatory reflexes but of respiratory reflexes as well and that the stimuli can be chemical as well as mechanical. "The intrapulmonary or humoral accumulation of carbon dioxide in the trunk constitutes the vagal reflex stimulus of the respiratory center of the isolated head."

From their experiments Heymans and Heymans concluded that the respiratory reflexogenic zones which are affected by chemical stimuli are found in the central organs of the circulation, the heart and the aorta, and that the lungs have no specific excitability to carbon dioxide. In our judgment this last proposition has not been confirmed. In the first place it should be remembered that it is more or less asphyctic blood circulating through the lungs which is used as a stimulus and *not the inspired air*, which would affect the whole respiratory tree, from the nose to the pulmonary alveoli. Thus it is rather the functions of a reflexogenic zone in the circulatory apparatus which are being investigated—the sensory innervation of the pulmonary vessels *stimulated by changes in the composition of the blood*—and not whether the respiratory pathways respond to the composition of the air. Again, the important operative intervention, extirpation of the heart and great vessels, may suffice to put the animal in an unphysiological state; and one must reckon also with the possibility that many sensory fibres of the pulmonary plexuses which pass to the vagi may be damaged by the same operation. All of this may damage or alter the vagal sensitivity and cause the vagotonic type of respira-

tion observed and the scarcely appreciable reaction to the circulation of the hypercapnic blood. It is invalid to argue that the Hering-Breuer reflex, which is known to be very rough, is still present. Anrep and Samaan (1933) have shown in denervation experiments that a very few remaining fibres are sufficient to cause this reflex to persist.

After the demonstration by Heymans and Heymans of the existence of circulatory reflexogenic zones with chemo-receptors we decided to return to the subject of pulmonary sensitivity. In new publications with Puche (1930) and with Raventós (1931–33) we employed the technique of the isolated head as usual, and in order to exclude all circulatory influence we bled the trunk "B" completely and rapidly by cutting the abdominal aorta until the heart stopped. Thus all circulatory factors both mechanical and chemical are suppressed since the blood no longer circulated. It is certain that with this bleeding the sensitivity of the lungs is diminished; but this fact, which furnishes a valid objection to negative interpretations, is favorable the other way around; in the normal state the reflex would logically be more intense and more effective.

With the addition of CO_2 in varying proportions to the air going into the trunk, while identical mechanical conditions of respiration are maintained, it is shown that even in such unfavorable circumstances the inhalation of CO_2 intensifies the movements of the isolated head which are already increased as a result of hemorrhage and subsequent asphyxia of the tissues.

We have tested also the response to the inhalation of irritant gases, hydrogen chloride and ammonia, and confirmed the observations made with Bellido (1919).

Heymans has (1929–33) raised the objection to the conclusions from these experiments that in the production of the reflexes observed it is not a specific sensitivity which is acting, but the general sensitivity to irritant agents. "The results obtained by Pi-Suñer and Bellido," state Cordier and Heymans (1935) "are certainly due to the fact that these authors have administered by inhalation air with concentrations of CO_2 which pass beyond physiological limits and even beyond the pathological. It is a question of phenomena of pharmacological order." It is difficult to know where a specific chemical excitation ends and where an irritant, nociceptive excitation of chemical origin begins, and where is located the boundary which separates physiology from pharmacology. HCl and NH_4OH for example do not increase the depth and rate of respiration like CO_2, but they inhibit it or provoke abnormal reactions: cough, spasm, etc. Barcroft and Margaria (1932) observed that high concentrations of CO_2 cause a characteristic type of respiration, which is not a simple increase but may be faster or slower depending upon the previous rate. After section of the vagi the breathing of high concentrations of CO_2 in air is usually followed by completely abnormal symptoms analogous to those observed by Lumsden (1923) upon sectioning the medulla at various levels, and similar to those which Taylor (1930) observed after the administration of cyanide: apneusis, gasps and finally paralysis. Some of our positive results have been obtained with mixtures of air and CO_2 which are closer to the normal and in which it cannot be said that carbon dioxide acts as an irritant gas.

The negative results of Partridge (1933), who attempted to record vagal action currents using inspired CO_2 as a stimulus, have also been cited in opposition to the theory of pulmonary chemical sensitivity. This is a problem which we began to study with Bellido in 1921. The electrovagograms which we obtained then showed certain differences when the animal breathed carbon dioxide, but were not conclusive. Our technique at that time was deficient. Adrian (1933) states that concentrations of CO_2 in air higher than ten per cent cause a modification, a slight diminution in the intensity of the impulses passing in the vagus when the lung is distended, but that these effects are small and within the limits of possible experimental error. Bullring and Whitteridge (1943–44) say that an increase in vagal electric discharge—single fibre preparation—occurs as soon as a volatile anesthetic—ethyl chloride, chloroform, ether, divinyl ether, trichlorethylene, etc.—reaches the lung, whether it is administered by inhalation or intravenously. But it is not possible to state that the respiration of air with carbon dioxide in greater concentration than normal is accompanied by definite signs in the electrovagogram.

In a later series of experiments (Pi-Suñer and Raventós, 1931) we have returned to the question, making the conditions of observation more rigorous even at the cost of simplicity. Using the method of the isolated head, we placed a cannula in the pulmonary artery and another in the left auricle of the trunk of "B." All of the rest of the heart and the thoracic aorta were removed, care being taken to avoid damage insofar as possible to the nerve fibres which form the plexuses at the roots of the lungs and their vagal continuation. Defibrinated, oxygenated, blood was perfused through the lungs by means of a Dale Schuster pump. Thus only the lungs and head of dog "B" remained alive, connected by the vagi which had been cut below the heart. After this procedure it is still possible to obtain respiratory responses to the inhalation of air with CO_2. It is evident that such reflexes cannot originate in the heart and aorta which are no longer present, and that they are due to the influence of the inspired carbon dioxide upon lungs which are perfused with blood of constant gas content.

Another series of experiments recently completed (1938–42) consisted in denervating the heart and great vessels while leaving the pulmonary innervation intact insofar as possible. Again we used the technique of the isolated head perfused with blood from a donor dog "A." Relying upon the anatomical findings of Lim (1893), Cannon, Lewis and Britton (1926) and Barry (1935), we established a technique of intrathoracic denervation of the heart similar to that described by Anrep and Samaan (1933) and Anrep, Pascual and Rossler (1932).

With this procedure the breathing of air with CO_2 by the trunk gives rise to changes in the respiratory movements of the head. Now the responses to chemical stimuli cannot arise in the heart and aorta which have been denervated, nor from the carotid sinus which, as in all the experiments with the isolated head, remains in the head and is perfused by blood from the donor dog.

It is important to mention the work of Dirken and Van Dishoeck (1937) who also demonstrated the sensitivity of the lung to changes in the concentration of CO_2 in the alveolar air and concluded that the nerve endings present in the lungs

are sensitive to changes in the concentration of CO_2 such as are found normally in alveolar air, i.e., 5 to 6 per cent.

Dirken and Van Dishoeck consider furthermore that the sensitivity of the pulmonary receptors is limited to CO_2 because they did not note any difference in the reflex respiratory responses, when the proportions of oxygen and nitrogen were varied in the breathed gas mixtures. Bilateral section of the vagi abolished the reflex effects of changes of CO_2 in the air.

After confirming the sensitivity to CO_2 of pulmonary and bronchial chemoreceptors, the authors reach a paradoxical conclusion: these receptors are not the instruments of the normal regulation of the respiratory movements. They observed in some experiments that as the frequency of the inspirations increased their amplitude was reduced, and they inferred from this fact that the ventilation was not changed. They doubt that the phenomena observed depend upon a direct sensitivity to carbon dioxide on the part of the pulmonary endings, attributing them rather to changes in the excitability of the receptors for their natural physiological stimulus: the collapse and the expansion of the lung evoking reflexes through mechanical stimuli. However the fundamental fact is the existence of endings which are sensitive to changes in the concentration of CO_2 in the alveolar and bronchial air and which influence the movements of respiration.

Recently Hammouda, Samaan and Wilson (1942–43) reaffirmed once again that the fibres in the vagi which conduct the afferent impulses of the reflexes of inflation and deflation are of pulmonary origin, and Bagoury and Samaan (1941) have studied respiratory reflexes by injecting ketone-bodies into the pulmonary circulation. The authors prevented any possible central action of the blood; the vagal excitation arose from sensory endings of the lungs. These pulmonary endings would therefore be sensitive both to mechanical and to chemical stimuli.

Eppinger, Papp and Schwartz (1924) infer the mediation of pulmonary chemical sensitivity in studying the mechanism of cardiac asthma; Dunn (1920) and Binger and Moore (1927) consider also local pulmonary influences when they attempt to interpret the hyperpnea of pulmonary embolism, and Churchill and Cope (1929) do likewise in considering the dyspnea observed in pulmonary edema and which they attribute to the excitation of sensory endings in the lung.

Another important aspect is the debated intervention of the lung in the maintenance of vagal tone. Heymans and Heymans observe a disappearance of the respiratory tone, maintained by the vagus, when the stimuli arising in the heart and aorta are no longer present and only those arising in the lungs remain. Heymans and Heymans conclude that the lungs are not the site of origin of the centripetal vagal impulses which maintain the physiological respiratory vagal tone and the reflex excitation and inhibition of the respiratory center in relation to respiratory and circulatory mechanisms.

Our observations demonstrated that after bleeding the trunk and therefore excluding all circulatory influence either mechanical or chemical, the normal rhythm of the respiratory movements of the head continues. *It is only after section of the vagi that respiration of the post-vagotonic type appears.* This suggests

that at least the sensory innervation of the lung assists in maintaining the vagal tone.

The latter proposition is confirmed by the investigations of Anrep and Samaan (1933). These authors cite the work of Pavlov (1895–96) and his pupils and collaborators, Cachovsky (1899) and Cheshcov (1902), studying the respiratory movements in vagotomized dogs surviving up to nineteen months after operation. Just as in the acute experiments, the rate is reduced to 4–8 per minute immediately after the section of the vagi, speeding up in the second or third week, and slowing down again to the same original bardypneic rate which is maintained until the death of the animal. Pavlov attributed the initial slowing to the cutting of sensory fibres of pulmonary origin and the subsequent acceleration to a process of irritation in the same fibres with scarring of the central stump of the cut nerve. When the process of cicatrization is completed the rhythm characteristic of the absence of vagal afferent impulses is re-established. Sharpey-Schafer (1932) states that the respiratory slowing is due to the increased resistance to the passage of air through the respiratory pathways owing to paralysis of the laryngeal nerves. It has not been difficult to demonstrate the error of this opinion. Heymans and Heymans (1927) assert that the post-vagotonic rhythm is due to the lack of cardio-aortic control. The "respiratory tone" is of cardio-aortic origin and disappears when the vagal fibres arising in that region are cut. "No direct evidence in support of this view is, however, provided," write Anrep and Samaan, who have shown that section of the vagal fibres at the level of the root of the lung does not alter the cardio-aortic innervation and that the branches which leave below this point go to the lung, with the exception of a few which pass to the esophagus. Taking into account these anatomical findings, Anrep and Samaan (1933) denervated the heart and aorta by means of intrathoracic dissection without producing the respiratory rhythm typical of vagal section. The section of one of the vagi and cardio-aortic denervation on the opposite side did not cause vagotonic respiration either. Later section of the remaining vagus in the neck caused this type of respiration to appear at once. Moreover, they sectioned the vagus below or at the level where the pulmonary branches leave the main trunk and compared the different results in the two cases. With the low section no differences in rate were observed; with the high section the characteristic slow respiration appeared; it switched from 24 to 7 respirations per minute. Subsequent bilateral vagotomy in the neck no longer modified this slow rate. This observation produces additional evidence against the cardio-aortic theory as well as against the laryngeal theory of Sharpey-Schafer.

Bouckaert and Heymans (1933) comment: "The reflex and tonic influences of the pulmonary vagus on the activity of the respiratory centre is a well-known and generally accepted fact (Hering and Breuer, 1868; Head, 1889; Haldane, 1922; J. F. and C. Heymans, 1926; C. Heymans, 1928; Hoffmann and Keller, 1929; Hess, 1931; Anrep and Samaan, 1933). It has been shown by J. F. and C. Heymans (1926) and C. Heymans (1928, 1929a, b) that the vago-depressor nerves in dogs are also the centripetal paths of respiratory reflexes in relation with the

cardio-aortic blood-pressure . . . . The normal arterial pressure in the left heart and aortic arch maintains a respiratory reflex tonus. Anrep and Samaan (1933) cannot find any evidence that cardio-aortic impulses invariably exert on the rate of the respiration a regulating influence which can be compared with the definite and constant dependence of the respiration rate on the pulmonary innervation.

Recently Sidney Harris (1945) has studied the behavior of the respiratory tonus in anoxia. During the reduction of oxygen content in inspired air to about 8 per cent the volume of the cat's chest at the end of expiration increased about three times the volume of a normal respiration. After vagotomy the anoxic increase in expiratory volume was about one-third as great as with the vagi intact. Crushing the nerves from the carotid chemo-receptors had little or no effect upon the reaction. An excess of the CO_2 in the oxygen deficient air did not prevent the increase in respiratory chest volume, but it did greatly increase the minute-volume of respiration. Gesell and Moyer (1934–35) and Green and Swanson (1938) had earlier reached similar conclusions.

Hermann, Jourdan and Vial (1934) suggest as a result of their experimental observations that the tone of the pulmonary vagus is continuous, depressing the activity of the cardio-inhibitory center and eliciting circulatory pressor reflexes which are antagonistic to the depressor reflexes arising in the cardio-aortic sensory area and in the sensory area of the carotid sinus. These conclusions, perhaps too schematic, are nevertheless of interest because they are based on experiments which demonstrate once again the rôle of the pulmonary vagus in controlling the respiratory tonus.

The existence of sensory endings in the respiratory apparatus should now be considered. Schumacher (1902), Hudovernig (1907), Tello (1924), Perman (1924), Gaylor (1934), demonstrated these receptors in the alveolar duct and the presence of fibres of pulmonary origin in the vagus nerve. Larsell (1921–39), Larsell and Burget (1924) and Sunder-Plassmann (1933) describe sensory nerve endings in the mucosa of the bronchi and trachea which are probably sensitive to mechanical stimuli. According to Larsell (1939), receptors of a different morphological type are found in the depths of the respiratory tree and it may be inferred that they are chemo-receptors because of their location directly in the air passage.

The fibres proceeding from all these endings ascend with the vagus, but Beccari (1934) believes that some afferent respiratory fibres also pass with the sympathetic, a fact in accord with the scheme of double sensory innervation of the vegetative organs through the sympathetic and parasympathetic outlined by Pi-Suñer and Puche (1928–30). Brookhart and Steffensen (1936) observe respiratory effects when the stellated ganglion is ablated, but they do not conclude that the afferent fibres implicated in the Hering and Breuer's reflex should necessarily pass through it.

Numerous facts, established by us and by other authors, demonstrate the chemical sensitivity of pulmonary origin. Heymans and Heymans have shown that the cardio-aortic areas are sensitive to chemical stimuli, and they think that the latter sensitivity is linked to the circulatory apparatus and excludes the

pulmonary chemical one. We consider, however, that no incompatibility exists; that, on the contrary, the effects of peripheral excitation of separate origin support each other functionally.

II. It is a well established fact that different sensory regions operate synergistically in the control of distinct vegetative functions. In the case of the respiratory and circulatory reflexes the sensory regions are various and do not hinder the function of one another but rather coordinate their regulatory influences.

Since the early observations of Pagano (1900) and Siciliano (1900) and the description by Hering (1927–32) of the functions of the carotid sinus, a number of authors have devoted themselves to this very important reflexogénic zone. The sinus is concerned not only with the regulation of circulatory dynamics under the influence of changes in the blood pressure, but its action extends also to various functions; among these is the respiratory.

"Many clinicians," write Heymans, Bouckaert and Regniers (1933), "have called attention for some time to the fact that carotid compression elicits changes in respiration." Tschermack (1866), Quincke (1875), Sollmann and Brown (1912), Danielopolu, Asland, Marcou, Proca and Manescu (1927), Danielopolu, Manescu and Proca (1928), Wenckebach and Winterberg (1927) proved that traction on the cephalic stump of the recently cut carotid awakes reflexes through mechanical excitation of the sinus. Similar findings have since been demonstrated repeatedly: Gollwitzer-Meyer and Schulte (1931), Schmidt (1932), Winder, Winder and Gesell (1933), Gemmill and Reeves (1933).

Moisejeff (1927) demonstrated that such effects originate in the sinus. Upon ligating the vessels which arise from the bifurcation of the carotid and perfusing the sinus he found inhibition of the respiratory movements when the pressure of the perfusing fluid increased. C. Heymans (1929) and Heymans and Bouckaert (1929–30) have confirmed these results. Afferent nerve impulses arise in the carotid sinus as a result of mechanical excitation and intervene in the process of regulation of the respiratory dynamics just as in the case of the circulatory. Houssay and Orias (1934) have studied the effects of excitation of the sinus on the contraction of the smooth muscle of the bronchioles and have observed inconstant reflex effects.

But not only do the mechanical conditions of the circulation influence the respiration through carotid and aortic body reflexes. *The composition of the blood is equally effective*, particularly the content of CO_2 and oxygen.

Heymans, Bouckaert and Dautrebande (1930–31–32) using the previously described technique of Moisejeff, discovered that acapnia of the blood produces reflex inhibition of the respiratory center while blood with an excess of CO_2 stimulates the center. According to Heymans, Bouckaert and Regniers (1933) the most effective chemical stimuli for the carotid body are hypercapnia, anoxemia and increased concentration of hydrogen ions. There immediately appeared numerous confirmations: Owen and Gesell (1931), Schmidt (1932), Selladurai and Wright (1932), Bernthal (1934), Heymans, Bouckaert and Samaan (1935), Gayet, Bennati and Quivy (1935), Samaan and Stella (1935), Zottermann (1935), etc. Slight changes in the composition of the blood passing through the

carotid sinus give rise to vigorous circulatory and respiratory responses. Hill and Flack (1908), Piras (1922), Frey (1923) and Hess (1931) did not believe that the respiration could be influenced normally in reflex fashion by a greater or lesser concentration of CO_2 in the blood. Hering (1932) and Mies (1932) discuss the chemical sensitivity of the carotid body; Gollwitzer-Meyer and Schulte (1931) and Gollwitzer-Meyer (1934) the specificity of the sensitivity to CO_2. Such a mass of confirmatory results was so gathered together that there is no possible doubt.

Anatomical investigation has also been here of great value. De Castro (1926-28) began the studies by describing a complex system of receptors in the carotid sinus. He explored immediately the carotid body or glomus at the bifurcation of the carotids in anatomical continuity with the sinus. It is a small organ which has the appearance of an endocrine gland, possesses certainly an intense metabolism and among its cells is distributed a profuse net of nerve endings. Sunder-Plassmann (1930) furnishes interesting data and Heymans, Bouckaert and Regniers (1933) give in their turn a description of the region. Ask-Upmark (1935) has studied the comparative anatomy of the sinus in twenty-seven species of animals. Nonidez (1935–36) and Boyd (1937) have found glomerular tissue in other vaso-sensory regions: principally in the aorta (Penitschka, 1931)—"Glomus aorticum"—and in the vicinity of the pulmonary artery—"glomus pulmonale." There are other analogous cellular formations, e.g., the coccygeal body. And perhaps one might refer to this group of small organs the groups of cells attached to or near the peripheral vessels described by Goormaghtigh (1935) which he considers to belong to the neuro-vegetative system. The glomera are probably a part of the chromaffin system even though their cells do not usually show the characteristic protoplasmic granules which are stained by chromates. The glomera seem to respond especially to chemical stimuli, above all those which are found in asphyxia: anoxemia, hypercapnia, acidosis.

In 1931 Heymans, Bouckaert and Dautrebande observed that ligature of the nerve fibres arising in the carotid sinus prevents reflex adaptation to changes in the pressure of the blood contained in the sinus but does not abolish the reflexes elicited by chemical stimuli. And reciprocally according to Heymans and Bouckaert (1932) it is possible to inactivate the chemo-receptors by producing emboli with lycopodium powder in the vessels of the glomus while preserving the sensitivity to pressure. Gollwitzer-Meyer and Schulte (1931) and Gollwitzer-Meyer (1934) observed that lobeline exerts its action, which is identical with that of anoxemia, only when it reaches the carotid body. Danielopolu, Asland and Marcou (1933) also separate the location of presso-receptors of the sinus and chemo-receptors which are in the carotid bodies. Camus, Bernard and Merklen (1934) brought new confirmation by cutting the fibres from the presso-receptors, and Comroe and Schmidt (1938) conclude that the receptors for pressure are found particularly if not exclusively in the carotid sinus itself, while the chemo-receptors are at the origin of the occipital artery near the carotid body and probably in the carotid body itself.

The study of the electrical variations produced in the nerve of Hering due to

the functional state of the carotid sinus has given data which demonstrate also the topographical separation of the receptors. Bronk and Stella (1932) observed that action currents in the nerve result not only from distention of the sinus by increase in the blood pressure, but can be produced also by other kinds of stimulation. Heymans and Rijlant (1933) showed the presence of impulses in the same nerve which bore no relation to the pressure in the sinus and were dependent on the state of ventilation of the animal: when the ventilation is poor the discharges are more intense; the excitation which sets up the impulses is not mechanical but chemical depending upon the gas content of the blood. Bogue and Stella (1934–35), Zottermann (1935), Euler, Liljestrand and Zottermann (1939) confirm these observations in experiments on cats and suppose that the starting point of the impulses is the carotid body. Bronk and Stella (1934) confirm the independence of the chemo-receptors and the presso-receptors. Samaan and Stella (1935) have studied the influence of changes in the composition of the blood on the action currents in the nerve from the carotid body and have shown that these currents cease when the tension of carbon dioxide in the blood is reduced below 32–35 mm. Hg. The physiological tension of carbon dioxide in the arterial blood will maintain therefore a tonic excitation of the carotid chemo-receptors and through them of the respiratory centers as has been shown by Selladurai and Wright (1932) and by Witt, Katz and Kohn (1934). Euler, Liljestrand and Zottermann (1939) agree with the earlier conclusions but consider that the physiological tension of oxygen is equally as effective as that of carbon dioxide.

In the aorta the two types of receptors, mechanical and chemical, are likewise found separated from one another. Comroe (1939) has studied the aortic chemoreceptors and particularly the aortic body which is similar to the carotid body, and has established that the bodies are in both cases the point of origin of circulatory and respiratory responses to anoxia; in the dog the aortic receptors are said to elicit reflexes which are predominantly circulatory while the carotid body is the principal point of origin of respiratory reflexes. There are nevertheless individual differences, and in the cat the functional specialization does not seem to be so clear. Gellhorn and Lambert (1939) believe that the circulation is regulated by the presso-receptors while the respiration depends principally on the chemo-receptors.

III. With the demonstration of the chemical sensitivity of the various vascular zones, the most important of which are the carotid sinus and the aorta, it was logical to study the physiological significance of this sensitivity in comparison with that of the respiratory centers. "Central chemical control," writes Gesell (1939) "which had been accepted on faith, was not on the defensive. In the confusion, many adopted new faiths which carried them too far, for luck was with the majority. This they do with the same enthusiasm as their predecessors who thought that the respiratory regulation and also the circulatory was carried out exclusively by the state of the blood supplying the centers." Schmidt and Comroe (1941) state: "For many years the physico-chemical aspects of respiratory control were dominant, but during the past decade the reflex factor has grown

steadily in stature until, in our opinion, the time has come to reverse the traditional attitude".

In making a comparison between the central and the peripheral effect of the chemo-receptors Heymans considered that, at least in the case of anoxemia, the effect of excitation of the sinus takes precedence over that of the centers in the control of the movements of respiration. Heymans, Bouckaert and Dautrebande (1930) observed that as a result of denervation of the vaso-sensory zones, only a moderate acceleration of the rate of respiration and a slight increase in the blood pressure is produced when the dog breathes nitrogen without oxygen, instead of the response of violent hyperpnea and extreme hypertension by the normal animal.

A number of investigators have obtained similar results. Selladurai and Wright (1932), Witt, Katz and Kohn (1934), Euler and Liljestrand (1936), Euler (1938), and Gesell and Lapides (1938) observed that in general respiration is depressed after simple denervation of the carotid sinus. Schmidt (1932) and Gemmill and Reeves (1933), studying the immediate effects of denervation of the sinus, had shown that anoxia does not produce an increase in the pulmonary ventilation, but rather a decrease. Schmidt and Comroe (1940–41) and Moyer and Beecher (1942) agree in the statement that hypoxia increases the respiration of animals whose vaso-sensitive zones have been deafferented. Nevertheless these animals have gone through a long period of depression. Watt, Dumke and Comroe (1943) assert also that anoxia produces stimulation of the vaso-chemoreceptors while it leads primarily to depression of the centers. In experiments in which the denervation had been carried out previously, performed when the animal had recovered from the operation ("chronic" experiments), Gemmill, Geiling and Reeves (1934) observed a slight excitation produced by anoxia, but only when it was not excessive. Wright (1934) in "acute" experiments on anesthetized rabbits and later in chronic experiments confirms the respiratory depression caused by anoxia after carotid denervation. Smyth (1936–37) states that this depression is considerable and, moreover, anoxia reduced the intensity of the response to CO_2 in denervated animals. In the rabbit Stella (1935) does not find significant effects of denervation: the respiration may decrease in depth but increase in rate. Schmidt, Dumke and Dripps (1939–40) have blocked the sinus nerves with procaine and observed scarcely any change from the normal respiration.

Wright (1936) remarks that after denervation of the sinus and bilateral section of the vagi, oxygen-lack does not produce hyperpnea. After the sensory nerves of the carotid sinus and the aorta have been removed Wright finds that severe anoxemia in anesthetized cats instead of causing respiratory activation produces a reduction in the ventilation. The changes in the respiratory gases in the blood are greater in the operated animals than in the normal. He concludes that up to a certain point the vascular receptors in the carotid body and the cardio-aortic area protect the centers from chemical influences "in loco."

Schmidt (1932), Beyne, Gautrelet and Halpern (1933); Winder (1933);

Mulinos (1934); Henderson and Greenbarg (1934); Wright (1934–36–38) Gayet, Bennati and Quivy (1935); Brewer (1937), Gesell and Moyer (1937); Smyth (1937); Comroe and Schmidt (1938); Gellhorn and Lambert (1939); etc., agree that anoxemia is the most effective stimulus for the vascular chemoreceptors, particularly of the carotid body.

Bernthal (1938) by perfusing the isolated carotid sinuses with blood of fixed pH and varying the concentration of oxygen, observed that 18 per cent oxygen causes at the most a moderate vaso-dilatation and hypopnea, while 15 per cent oxygen gives rise to vaso-constriction and hyperpnea an effect intensified when the proportion of oxygen is reduced to 12.8 per cent. A drop to an oxygen tension of 10 mm. Hg causes a marked increase in vaso-constriction and respiratory rate.

In spite of these facts, Cromer and Ivy (1931) noted that dogs from which the carotid sinuses had been removed aseptically, work in the treadmill without greater effort than normal dogs and do not show any respiratory disturbances. It should be noted however that in these dogs the cardio-aortic receptors remained intact. Decharneux (1934) also observed adequate respiratory responses afte denervation of the sinus. Dautrebande (1937) stated that dogs with complete deafferentation of the sinus and aorta respond to low atmospheric pressures in the same regular and adequate fashion as do normal dogs. Gesell (1939) cites cases of hyperpnea in response to anoxemia after denervation of the sinus.

It has long been known that the cyanides and sulfides inhibit oxidative processes in the organism. The mechanism of action of cyanides is well-known. Cyanide inhibits cellular oxidations by combining with cytochrome oxidase. Sulfides act in the same manner. A. Pi-Suñer and J. Pi-Suñer (1928–29) affirm that to the extent that the cyanides impede the processes of oxidation, in particular of glucose and fat, they evoke trophic reflexes which produce hyperglycemia and hyperlipemia.

Haggard and Henderson (1922) have studied the effects on the respiratory movements of sodium sulfide injected intravenously. Heymans and Heymans (1927) and Heymans, Bouckaert and Dautrebande (1930) showed, and it was quickly confirmed by Owen and Gesell (1931), that the injection of sulfide or cyanide into the carotid sinus causes very intense respiratory responses. Winder and Winder (1933) examined the action of the sulfides and called attention once more to the strong respiratory reflex effect elicited from the chemo-receptors of the carotid body. Denervation of the sinus modifies the response: doses which were previously extremely effective evoke, after denervation, a very different type of response of slight intensity. In all cases the reactions to stimulation of the sinus are more intense than those obtained by direct application to the centers. Winder, Winder and Gesell (1933) required quantities 25 to 75 times as great after denervation of the sinus to produce occasional respiratory effects. Winder's (1937) statement that monoiodo acetic acid acts upon the carotid body in a similar way as cyanides and sulfides is not in agreement with the general opinion that the first is an inhibitor of sulfhydryl enzymes while cyanide inhibits cytochrome oxidase.

These drugs also act upon the respiratory centers. Winder, Winder and Gesell (1933) applied cyanide directly to the fourth ventricle but noted that the respiratory movements were simulated only by larger doses than those required on the sinus; further increase in the doses caused difficulty of respiration which finally ceased. Injection of sulfides and cyanides into vertebral artery requires larger doses in order to alter the respiratory movements than are required with injection into the sinus.

Another factor which should be kept in mind in considering the action of cyanide or sulfide is the rôle of the vagus. According to Haggard and Henderson (1922) the respiratory effects in these cases are due to the action of the drugs on the vagal endings of the lung, in view of the differences in the responses depending on whether one is dealing with intact or with vagotomized animals.

Winder (1937), Bernthal and Weeks (1939) arrived at the conclusion that changes in pH in the blood and in the sensitive cells are a prime factor in normal carotid body stimulation, and Von Euler, Liljestrand and Zottermann (1939–41) confirm that intracellular acidity in the carotid body is increased during anoxia.

Dripps and Comroe (1944) divide the drugs that affect the sensibility of the carotid body in two groups: those that, like cyanides and sulfides, produce effects by inhibiting intracellular respiratory enzymes, and those which, like lobeline, are synaptotropic and consequently affect the transmission of afferent impulses. Hollinshead and Sawyer (1945) conclude from the above facts that a chemical agent mediates in the excitation of the carotid body. They suggest that this agent would not be acetylcholine.

Not only are there differences in intensity between the reflex effects of anoxemia and those of central origin, but there are also differences in the speed of the response. This is shown by the frequently cited observations of Heymans, Bouckaert and Dautrebande (1930), of Gemmill, Geiling and Reeves (1934), and of Henderson and Greenberg (1934). These authors observed that the respiration is not increased during the first 4 to 60 seconds of asphyxia due to oxygen-lack produced by breathing pure nitrogen after denervation of the carotid sinus; but that if the asphyxia is continued the hyperpneic reaction appears in the second minute. Carbon dioxide contained in the blood also stimulates the vascular chemo-receptors. In this case however no dominance of the chemo-receptors over the centers is observed. Heymans, Bouckaert and Dautrebande (1930) were able to show that the respiratory response to the breathing of air with CO_2 is produced in the same manner after denervation of the carotid sinus as in the normal state. Gemmill and Reeves (1933); Stella (1935); Wright (1934–36); Gesell and Moyer (1937) demonstrated also intense respiratory reactions to CO_2 after denervation. On the contrary Selladurai and Wright (1932); Schmidt (1932); Green and De Groat (1935); Euler and Liljestrand (1936) observed that with denervation the responses to CO_2, like those to anoxemia, are reduced although never to such a great extent.

It is probable that the differences in the results are due to various factors, the species of animal and the anesthesia. Marshall and Rosenfeld (1937); Comroe and Schmidt (1938); Schmidt and Comroe (1940) believe that the reflexes play a

more important part in anesthetized animals than in those in the normal state precisely because of the effect of the anesthetic on the centers. And they add (1941): "We did not believe that chemo-receptor reflexes are an important factor in maintaining eupneic respiration or in bringing about the respiratory response to carbon dioxide under ordinary conditions." They believe that respiratory centers are more sensible to carbon dioxide than chemo-receptors are.

Gayet, Bennati and Quivy (1935) found intense hyperpnea when the carotid sinus was perfused with equal parts of blood and Locke's solution and the CO_2 tension was increased in the perfusion fluid, but they consider that the effect of CO_2 on the centers is stronger than on the chemo-receptors. Bernthal (1938) and Comroe and Schmidt (1938) state that the reflex effects of the concentration of CO_2 in the blood are less marked, less constant and less well maintained than those produced by anoxemia. Schmidt, Comroe and Dripps (1939) affirm that the threshold of carotid sensitivity to CO_2 is relatively high and always above the threshold of the centers, at least in vagotomized dogs under light anesthesia. Schmidt, Dumke and Dripps (1939–40) confirm this finding; Comroe (1939) points out the existence of marked individual differences in this respect.

Heymans, Bouckaert and Regniers (1933), Gesell (1939) and Schmidt, Comroe and Dripps (1939) agree that great increases in the concentration of hydrogen ions in the blood stimulate the chemo-receptors as strongly as anoxia and more than hypercapnia. Boycott and Haldane (1908) attributed the hyperpnea of anoxemia to an accumulation of lactic acid in the centers. Winder, Bernthal and Weeks (1938) observed an increase in the respiratory movements when the vessels supplying the carotid body were ligated, and after removal of the ligature the hyperpnea and the hypertension disappeared.

Bernthal (1938) perfused the carotid sinus and concluded that in the normal state the carotid chemo-receptors are the origin of tonic vasoconstrictor and respiratory reflexes of great sensitivity which are controlled by the tensions of oxygen and carbon dioxide in the blood. After this Bernthal and Weeks (1939) showed that cooling the blood perfusing the sinus depresses the respiratory and vascular reflexes while warming the blood increases them; they attribute these effects to variations in the acid-base equilibrium of the receptors related to their metabolism. Stadie, Austin and Robison (1927) have already demonstrated the increase in acidity in the tissues as a result of warming.

Schmidt, Comroe and Dripps (1939), and Schmidt, Dumke and Dripps (1939–40) have repeated the experiments of Bernthal using a saline perfusion fluid instead of blood, and carefully determining the concentration of CO_2 and O_2 as well as the pH of the solution. They confirm the results of Bernthal and Weeks but they attribute the effects of the temperature changes to modifications in the gas content of the blood. Winder (1942) perfused the carotid sinus with heparinized blood in Locke's solution which carried CO_2 and O_2 in different concentrations. The effects of hypoxia and of hypercapnia are similar. The chemoreceptors are considered as one of several probable sites for mutual facilitation of hypoxia and hypercapnia acting as stimuli of respiration. Marshall and Rosenfeld (1937) obtained prolonged excitation of the sino-aortic receptors upon in-

jecting pyruvic acid cyanohydrin, probably due to slow, prolonged liberation of cyanide.

The most effective agent for the excitation of the centers and of the chemo-receptors, as Gesell has maintained since 1925, would be a change in the concentration of hydrogen ions in the neurones themselves when the internal environment, the blood or interstitial fluid, becomes acid or particularly when changes in the metabolism flood the neurones with acid metabolites. Warming, ischemia of the sinus or local poisoning with cyanides or sulfides all act in this way. Winder (1942) agrees that intracellular concentration of hydrogen ions is a factor in the control of chemo-receptor activity. Gesell, Krueger, Gorham and Bernthal (1930) examined some time ago the circumstances which can change the concentration of hydrogen ions in the tissues in relation to the insufficient supply of oxygen and the state of their metabolism. They suggest that impaired oxidation leads to increased acidity and that the reverse may also be true.

Bernthal (1938) has shown that anoxia, cyanide, hypercapnia and lactic acid cause reflex peripheral vasoconstriction through their local action on the carotid body, and the vascular reactions are accompanied by respiratory reactions. There is an exact correspondence between the activity of the chemo-receptors and their acidity. Nevertheless the respiratory effects of changes of the hydrogen ion concentration in the carotid body should not make us overlook the sensibility of the respiratory centers to these changes. Moyer and Beecher (1942) assert that decreased oxidations within the center constitute an important factor of respiratory adjustments due to changes in the central pH. This was confirmed by Comroe (1943). He applied upon the respiratory medullary center minute amounts of CO_2-bicarbonate mixtures and observed marked respiratory responses according to the pH of the mixtures. Garcia Banús, Corman, Perlo and Popkin (1944) found that in anesthetized dogs deprived of their chemo-receptor reflexes by the denervation of the carotid sinus and section of both vagi nerves, the respiratory compensation by the centers may be efficient enough to maintain the pH of arterial blood constant at least within 0.01 even if changes of O_2 and CO_2 tension occur in the inspired air.

Gesell (1939) shows that changes in pH in the nerve cells act in the same sense and in the same way in the respiratory centers and in the chemo-receptors: both are sensitive to carbon dioxide and to oxygen lack and hence to externally caused variations in the hydrogen ion concentration, but they are specially sensitive to variations in hydrogen ion concentration of internal origin brought about by the cellular metabolism. As a result of this central and peripheral sensitivity to hydrogen ion concentration, an adjustment is established between the nutritive requirements, the respiration and the circulation.

A characteristic example of these mechanisms of co-ordination is what Hess (1917) has called the "nutritive reflex" which has local and general vascular effects, as Fleisch (1938) and Rein (1938) among others have proved. It is known that any increase in metabolic activity raises the cardiac output. The most effective of the stimuli which give rise to these reflexes of nutritive origin is the acidity of active tissues.

Besides, as Gibbs, Gibbs, Lenox and Nims (1943) wrote, CO_2 improves the oxygenation of the tissues when the O_2 tension in air is low, which enhances the importance of hydrogen ion concentration in cells and in fluids, assuring fine respiratory adjustments from centers to tissues.

Nielsen (1936) suggests in opposition to the idea that the ordinary stimulus of the chemo-receptor centers is the increased local acidity, that carbon dioxide acts by virtue of specific properties and not because it is an acid. CO_2 is said to be the normal stimulus for the respiration: a chemical stimulus produced without interruption in the organism, the principal functional mediator in maintaining the physiological tone of the most diverse functions, especially the circulation and the respiration.

IV. As a result of the simultaneous action on the centers and on the chemoreceptors and as a result of the differences in excitability of centers and receptors depending upon the chemical stimuli acting upon them, various explanations have been developed for the adaptation of the respiration to the changing physiological needs. It is a difficult problem and in certain respects, as it often happens, an imaginary one. The opinions of authors differ in attributing preponderance to the peripheral factors or to the central ones in each case and greater or less physiological significance to the one or the other.

Haldane and Priestley (1935) point out the exquisite sensitivity of the respiratory centers to CO_2, which has long been known. Jongbloed (1936) schematizes the respiratory control stating that carbon dioxide excites the centers directly, while low tensions of oxygen cause tonic excitation of the chemo-receptors, particularly those of the sinus, from which impulses arriving at the centers sensitize them by lowering their functional threshold. Nielsen (1936), Henderson (1938) consider similarly that variations in oxygen to the point of anoxemia control the central activity which is set in motion by the presence of CO_2 in the blood. The tension of oxygen present affects the local action of CO_2 in the respiratory centers.

According to Bernthal (1938) the sensitivity of the carotid body to CO_2 is more variable than its sensitivity to oxygen, and decreases in the presence of an excess of oxygen. The normal tension of CO_2 in the blood exerts a tonic action on the function of the respiratory and circulatory centers by way of the chemoreceptors. Schmidt (1932) stated that the participation of the reflexes from the sinus in maintaining the respiratory movements is to be inferred from the marked reduction or abolition of the respiratory response to anoxemia after denervation of the sinus. But the respiratory reflexes of the sinus are essential only in the case of hyperpnea due to anoxia; apart from this defensive reaction, they do nothing which could not be done without them through changes in the blood flowing through the centers. Sensitivity of the cells of the center to the tension of CO_2 or to the concentration of hydrogen ions in the arterial blood is greater than the sensitivity of the carotid body. The nervous centers take precedence as the most highly specialized part of the mechanisms which control the adjustment of the respiration.

Schmidt insists once again (1941): "The part played by chemo-receptor reflexes in adjusting pulmonary ventilation to the requirements of the body must be negligible under normal conditions. This is indicated by the undoubted facts that the hyperpnea of exercise is not associated with any change in the blood that could stimulate these structures; and that the hyperpnea of carbon dioxide inhalation has not been found to be modified in any measurable way by removing the chemo-receptor influence.... The place of these reflexes in the body's economy lies, not in their sensitivity to factors which ... normally regulate breathing, but in their ability to withstand and to respond under adverse circumstances which seriously interfere with the functional capacity of the central neurons. By virtue of this ruggedness, the chemo-receptors are enabled to set up a powerful reflex drive and thus to maintain the activity of the neurons when the latter have lost their ability to respond to their normal stimuli.

Comroe and Schmidt (1938) consider also that the regulation is brought about by the tensions of CO_2 and oxygen in the blood bathing the centers and particularly by the concentration of hydrogen ions resulting from the equilibrium between the two gases. The chemo-receptors, which are more sensitive to oxygen-lack than to excess of CO_2, operate only when the deviations in one or the other or both of the two respiratory gases in the internal environment are very great and without sufficient effects on the centers.

Schmidt, Dumke and Dripps (1939) have made comparative studies of the sensitivity of the centers and of the carotid body to changes in the tension of CO_2 in the blood. The respiratory movements are not altered significantly nor does the tension of CO_2 in the arterial blood change after denervation of the carotid bodies so long as the animal is breathing sufficient oxygen. The carotid reflexes play an important part only in extraordinary circumstances and not concerned with the fine control of the respiration.

Gesell and Moyer (1937) made an experimental study of the central and peripheral factors which determine the frequency and depth of the respiratory movements. Lack of oxygen in the inspired air, and therefore in the alveolar air, causes a very obvious increase in the ventilation. This increase is not prevented by section of both vagi nor by section of the pulmonary fibres of the vagus and denervation of the carotid sinus, leaving the cardio-aortic region sensitive and active. Denervation of the sinus facilitates rather than prevents the acceleration, but it reduces the tidal volume. On the other hand denervation of the carotid and of the cardio-aortic region leaving the pulmonary vagus intact favors the acceleration produced by oxygen-lack. Complete denervation of the chemoreceptors of the sinus,the cardio-aortic area and the vagal pulmonary area eliminates the hyperpnea due to anoxia. It is evident that the lack of oxygen affects the chemo-receptors more than the centers. Hypercapnia acts in a different way: it causes a type of hyperpnea which is characterized principally by increase in the depth of respiration. The denervation of the vascular chemo-receptors, with or without vagal block, does not cause visible changes in this hyperpnea; this indicates that excess CO_2 exerts a stronger dominant stimulation upon the

centers than upon the peripheral receptors. From this equilibrium between the function of the centers and that of the receptors there results the adaptation to the various circumstances which can influence the respiration.

Schmidt and Comroe (1940) are still concerned with the functions of the carotid and aortic sensory zones. They reach the conclusion that in the case of anoxemia and of poisoning by cyanide or lobeline, the increase in ventilation is due to reflexes of vascular origin, from the sinus or the aorta. On the other hand, when an increase in CO_2 or in the concentration of hydrogen ions in the blood occurs, the central effect predominates. The reflex factor does not become active until almost the maximal response of the centers has occurred. They suppose that the different circluatory chemo-receptors show a distinct sensitivity to stimuli brought by the blood. Some of them might be active constantly in physiological conditions because of their greater sensitivity; the majority however function successively in increasing numbers only when the level of the stimuli is raised. A definite *quantum* of increase in stimulating activity is necessary to evoke a measurable reflex response corresponding to a definite threshold. The problem of whether the most diverse stimulating agents affect all the receptors or whether there is a specific sensitivity for each group of them has not yet been solved. The scheme proposed by Comroe and Schmidt is kept unchanged: the CO_2, the degree of acidity of the blood and perhaps of the nervous tissue, is the central stimulus and is almost without effect on the vascular chemo-receptors. Oxygen-lack on the other hand is the stimulus for the chemo-receptors which begin to function only in emergency. The same authors believe that the status of the hydrogen ion as a stimulus to centers and chemo-receptors is at present still uncertain. It seems quite possible that the respiratory effects of changes in pH (in so far as they are not referable to corresponding changes in CO_2 tension) are due to chemo-receptor reflexes and not to a direct effect of hydrogen ions on the centers.

Dumke, Schmidt and Chiodi (1941) try once more to elucidate the parts played by the peripheral chemo-receptors of the carotid sinus in the respiratory responses to anoxemia and to hypercapnia. With the animal breathing room air, quietly and without exerting any force, the ventilation rate is not decreased by denervation of the sinus, which demonstrates that in such conditions of repose the reflexes from the sinus play no appreciable part. Twelve per cent oxygen causes slight respiratory acceleration, which is increased when the proportion of oxygen in the gas mixture is reduced to 10 per cent. Denervation of the sinus makes the response to the drop in the inspired oxygen much less obvious. If 3.5 per cent carbon dioxide is added to the mixture containing 10 per cent oxygen, the respiration is increased; this increase is slightly less when the sinus is deafferented. From all this the authors conclude that hyperpnea produced by anoxia is reflex in nature, depending upon the stimulation of the vascular receptors, while the presence of an excess of CO_2 acts rather upon the respiratory centers. Gesell and Lapides (1938) performed an interesting experiment: they determined the duration of the apnea produced by over-ventilation of the alveoli in intact dogs and comparatively after blocking the nerve fibres of the sinus with cocaine.

With the nerves blocked the apnea lasts considerably longer. This means that with the afferent nerve of the sinus no longer conducting impulses and therefore with the chemo-receptors no longer functioning, a lower concentration of oxygen is required to excite the centers.

The observations of Witt, Katz and Kohn (1934) lead us to suppose that the participation of the chemo-receptors in maintaining the respiration is not dispensable. Deafferentation of the respiratory centers by section of the vagi and denervation of the carotid sinus results in depression of the respiration and in some cases in cessation and subsequent death. The impulses which normally reach the centers by the afferent nerves stimulate and condition the activity of the centers, and the respiratory mechanism depends to a great extent on the integrity of the peripheral receptors and its conductory nerves. Finally the authors consider that, in addition to the receptors of the sinus and to those which send their fibres through the vagus (pulmonary and cardio-aortic receptors) still others may play a part in the physiology of the respiration.

The centers and the chemo-receptors act in co-ordination in response to and according to the state of the gas content of the blood. Gesell (1939) asks whether there could be transitory and variable differences in the respective sensitivities and whether for that reason there might be produced physiological variations in control according to the circumstances prevailing in the respiration, which are always variable. Thus the threshold of a given chemical sensitivity would not be constant in a given place either central or peripheral.

If during intense hyperpnea by hypercapnia, the nerves of the carotid sinus are blocked with the vagi blocked as well, no decrease in the respiratory movements is observed at the time when the chemo-receptors should be stimulated most strongly. This could be explained in two ways: 1. The respiratory centers respond more vigorously than the receptors to high tensions of CO_2, while the reverse is true at lower tensions. 2. Hypercapnia is unable to maintain the reflexes in this case because the carbon dioxide in the centers blocks the impulses arriving from the chemo-receptors (Gesell and Moyer, 1935). Thus during intense hypercapnia, central stimulation of the vagus produces no effect on the respiration. It seems nevertheless that asphyxia ordinarily increases the effectiveness of afferent respiratory impulses and that the effects of chemical stimulation predominate in the respiratory centers (Gesell, Moyer and Kittrick, 1942).

There remain functional relations of great complexity between the peripheral receptors and the centers and variations would be possible in the relative participation of the mechanisms which control the functions of the centers and of the various chemo-receptors. Any rigid schema which is proposed in an attempt to explain these complicated phenomena can scarcely conform to all the facts.

V. There still remain other important factors which must be considered in studying the action of carbon dioxide on the centers and on the receptors in controlling the respiration. Not only is hypercapnia important, but also a factor of great influence is the lack of CO_2 sufficient to alter its necessary quantitative relationship with oxygen in the air and in the blood.

Mosso (1885) called attention to the effects of hypocapnia when rarefied air is breathed at high altitudes. Since then many have been carried out, some of which have had immediate practical importance. We refer particularly to the work of Yandell Henderson and his collaborators since 1908.

Hasselbalch (1912), Hasselbalch and Lindhard (1915) and Henderson and Haggard (1918) demonstrated that the ratio $\frac{H_2CO_3}{NaHCO_3}$ is altered by hyperpnea (L. J. Henderson, 1919). The concentration of H_2CO_3 in the blood is reduced by the increased expiration of CO_2 and the alkalosis must be compensated by a reduction in base, principally in fixed alkali, i.e., sodium, which is followed by an increased elimination of ammonia, sodium, etc., in the urine. On the other hand a higher concentration of carbon dioxide causes an increase of alkali in the blood. This has been confirmed repeatedly (Haldane, Kellas and Kennaway, 1919; Davies, Haldane and Kennaway, 1920; Gesell, 1922–23; Gollwitzer-Meyer, 1924; etc). Henderson (1925) and Gesell (1925) have summarized in two excellent reviews what was then known of the chemical regulation of the respiration, referring of course at that time only to the action of chemical stimuli on the centers.

When the respiratory elimination of CO_2 becomes excessive, as in the hyperpnea of anoxia for example, a paradoxical situation arises with increased pulmonary ventilation and simultaneous hypocapnia of the centers, in which the CO_2 is washed out. The more intense the hyperpnea of anoxia with subsequent hypocapnia particularly during anesthesia, the less will be the central control and the greater the peripheral control through the chemo-receptors. Hypocapnia reduces the excitability of the centers (Gesell and Lapides, 1938) and if it remains excessive the centers do not provide adequate ventilation. Anoxia results finally in respiratory depression, which is to be explained by the paralyzing effect, especially on the nervous functions, when the oxygen lacks (Verworn, 1903). The nerve cells are particularly sensitive, more so than any other cells, to anoxia and to asphyxia.

The chemo-receptors are in general strongly excited by degrees of anoxemia which depress or paralyze the physiological activity of the central neurones. But so long as the anoxia of the centers is not excessive the lack of oxygen increases their excitability to CO_2, the normal chemical stimulus. In this way the tension of oxygen in the blood would intervene in the control of the ventilation by its action on the respiratory centers. As the anoxia becomes more severe, asphyxia of the neurones hinders the process of regulation, impeding or even paralyzing the central responses; but at that time the chemo-receptors are still able to respond to the lack of oxygen in the blood and the respiration is maintained by the impulses which they set up as a compensatory reaction. This persistence of the excitability of the receptors leads one to suppose that the metabolic level of the sensory cells which compose them is lower than that of the central neurones.

Gesell (1939) holds to his old point of view: the regulation of the respiration is carried out by the acid-base equilibrium, by the pH in the nerve cell, and not by that in its environment, the interstitial fluid, still less by that in the blood. It

is the acid-base equilibrium of the centers primarily and also of the chemo-receptors. Only by taking into account the state of the acid-base equilibrium, not in the internal environment but in the intracellular fluid itself, is it possible to explain a number of facts which otherwise present insoluble problems.

Acidosis of the blood, for example, is not in all cases an obvious cause of hyperpnea. Henderson (1938) has shown that in carbon monoxide poisoning when there is an increase in the ventilation alkalosis is produced by hypocapnia which up to a certain point protects the centers from the harmful effects of anoxia.

We have seen the significance of the accumulation of acids in the cells of the centers and of the chemo-receptors; acids which arise in the intermediate metabolism of the cells and are also dependent upon the acids of the surrounding fluid, the internal environment. Gesell (1925) has explained how the respiratory centers function to maintain their optimal acid-base equilibrium. McGinty and Gesell (1925) have shown that the concentration of lactic acid in the brain and also in the blood increases when carbon monoxide is administered. The same thing happens in cyanide poisoning and in hemorrhage. Local accumulations of lactic acid and of other acids(pyruvic, citric, etc.) are possible in the nerve centers and other organs, in some cases perhaps without any increase in the acids of the blood.

We can explain now the danger of administering oxygen alone in attempting to combat anoxemia. The oxygen releases quantities of acid, particularly lactic acid, which disappears by oxidation and by resynthesis, and free base remains which binds CO_2, decreasing the normal stimulus for respiration. For this reason in respiration at high altitudes the addition of CO_2 to the inspired oxygen helps to increase the oxygen in the blood, to restore normal respiration and to produce the subjective improvement which is felt with each breath, as has been shown by Schneider, Truesdell and Clark (1926), and by many others, most recently by Dill (1938).

The final conclusion from all that has gone before is that for various reasons—lack of oxygen, excess of CO_2, hemorrhage, the action of certain poisons, and increased temperature in the centers or chemo-receptors—the general acid-base balance of certain organs can be changed and it is precisely the state of this balance which controls the respiration. One must not judge this equilibrium only from the pH of the blood; the local intracellular pH of neurones in the centers and of the sensory cells or neighboring cells in the peripheral chemo-receptors is of greater importance. Various factors act upon these equilibria and only by considering exactly the complex influence of such factors can one explain certain facts which otherwise would seem paradoxical if not contradictory.

Schmidt and Comroe (1940) report that many different substances evoke reflexes through their action on the chemo-receptors, particularly those of the carotid sinus. The active substances are of various sorts. We can explain the mechanism of action of some, according to what has already been said, by their influence on the nutrition of receptors through the cellular metabolism; this explanation cannot be applied to others to which we might attribute a

specific action, like that of so many drugs which act on various portions of the autonomic system.

It should be pointed out that the receptors are not only affected by variations in the hydrogen ion concentration, but that they may be excited by a variety of agents foremost among which are normal or accessory products of the intermediate or final metabolism of the tissues. It is possible that there is actually something more in these mechanisms than the regulation of oxidations and the production of acids. It may be that the processes of chemical excitation have greater significance and even though it has not been possible to confirm the mediation of endocrine products of the glomerular tissue in accordance with the hypothesis of Moniz de Bettencourt, Rodrigues Cardoso and Paes de Vasconcellos (1938), it would not be illogical to suppose that the capacity to stimulate the chemoreceptors and the centers extends to various substances of endogenous origin, metabolites or specific substances, "active substances" in the sense used by Demoor. In this way mechanisms of functional correlation and of regulation would be established, as a result of which the respiratory response could be perfectly adapted from moment to moment to the functional metabolic needs, which change rapidly according to the circumstances. Among those chemical stimuli which are able to evoke such mechanisms the most important is the concentration of hydrogen ions in the appropriate tissues.

VI. As long ago as 1876 Latschenberger and Deahna had formulated the hypothesis of the reflex influence of the peripheral circulation on the regulation of the respiration. Recently various investigators have sought by different methods to locate the vaso-sensory zones. Spalto and Consiglio (1886), Hager (1887), Pagano (1900), Siciliano (1900), Brodie and Russell (1900), Mayer, Magne and Plantefol (1920), Frey and Hagemann (1921), Haggard and Henderson (1922), Hess (1923), Tournade and Malméjac (1931), Tournade (1932), Tournade and Rochisani (1934), etc., demonstrated the sensitivity of the vessels and the effects of the chemical excitations which occur in them. The stimulation of these vascular receptors has effects which are predominantly circulatory, but which are inseparably linked to respiratory effects. Among these reflexes particular interest attaches to the respiratory reflexes which are produced by variations in the pulmonary circulation as a result of the stimulation of receptors in the vessels, as observed by Harrison, Harrison, Calhoun and Marsh (1932), Schwiegk (1935) and Christie (1938). Waele and Van de Velde (1940) describe respiratory reflexes through cardiac receptors.

In the same way receptors are found in various organs the excitation of which results in respiratory reflexes. Lewy (1891), Varaldi (1893), DuBois-Reymond and Katzenstein (1901–02), Baglioni (1903), Fleisch (1921), Scott, Gault and Kennedy (1922) have repeatedly confirmed the existence of such reflexes evoked by the excitation of sensory endings of the muscles. The most obvious effect is obtained from the diaphragm and from other respiratory muscles (Fleisch, 1928–30; Hess, 1931; Sharpey-Schafer, 1932; Gesell, 1935). Krogh and Lindhard (1917) produced respiratory reflexes upon faradic stimulation of various groups

of muscles. These effects may be similar to those which are observed when any sensory nerve capable of causing pain is stimulated.

The metabolic state of the tissues and particularly of the muscular system may therefore produce reflexes. It is well known that the respiratory activity increases automatically when the work is increased; this is due to central processes but still more to nutritive reflexes, reflexes which one may suppose are organized to constitute systems,—sometimes very extensive systems,—of respiratory reflexes. Because the increase in respiration begins at the same time if not before the commencement of work, Krogh and Lindhard (1920) thought that this was a matter of central effects, of efferent irradiation in connection with the muscular excitation from the centers (Paterson, 1928). Allen (1942) describes respiratory effects as responses to conditioned reflexes. All this presupposes the existence of reflexogenic receptors in the organs. Harrison, Harrison, Calhoun and Marsh (1932) and Harrison (1939) state as a result of their experiments on dogs and their observations of healthy and diseased human beings that reflexes evoked by movements of the legs help to produce the hyperpnea of exercise. Alam and Smirk (1937) showed the existence of a chemo-sensitive system in the muscles which is capable of evoking vasomotor reflexes as a result of the accumulation of metabolites produced in exercise. Comroe and Schmidt (1943) confirm the possibility of other similar respiratory reflexes resulting in hyperpnea and evoked by movements of the legs. These reflexes are thought to be elicited by excitation of the muscular chemo-receptors which are sensitive to inorganic metabolites, especially carbon dioxide. These mechanisms were suggested long ago by Volkmann (1841) and by Vierordt (1844).

Schmidt and Comroe (1941) and Comroe (1944) assert that in the hyperpnea of exercise the respiratory adjustment depends on receptions in the muscles, lungs (Christie, 1938; Harrison and others, 1939) and perhaps in the heart (Waelo and Van de Velde, 1940), besides many other well-known important factors.

The stimulation of chemo-receptors in the active tissues has been conclusively demonstrated as has their effect on the circulation and the respiration. We have referred above to the nutritive reflexes. Through their mediation the cardiac output arises as the metabolic exchange in the tissues is increased.

The regulation of the vegetative functions becomes progressively comprehensive until resulting in generalized responses. Thus, reversing the relationship stated above, respiratory stimuli affect the circulation in an entirely appropriate fashion. The addition of CO_2 to the inspired air increases the blood flow in the brain and decreases it in the muscular system as has been demonstrated repeatedly by Schmidt (1928) in the cat, Lennox and Gibbs (1932) in man, Irving and Welch (1935) in the rabbit, and Irving (1939) in the beaver and the muskrat. Bernthal (1934) observed that when the carotid sinus was perfused with blood containing oxygen at a tension 10 mm. Hg below the normal there is a reduction in the blood flow through the axillary artery. An increase in the CO_2 tension of the blood increases the flow. Section of the vagi makes these effects more obvious.

From all this it is to be concluded that the respiration like the circulation is controlled from the organs themselves by the presence of chemical stimulating substances. We have shown (1941) that a reflex increase in the respiratory movements is produced by asphyxia of the trunk when it has been isolated from the head and separated therefore from the carotid sinus and when the heart has been denervated, inactivating the cardio-aortic chemo-receptors. This effect can only be the result of the stimulation through CO_2 of receptors located in the lungs or in the tissues or in both regions.

VII. Afferent impulses of the most diverse origin produced by the excitation of receptors of all sorts reach the respiratory centers. Besides stimulation of mechanic and chemo-receptors, the stimulation of other more or less specific receptors is important. In the first place we may note the stimulation of many sensory nerves which evoke respiratory reflexes. Painful sensations can retard or accelerate the respiratory movements depending upon the circumstances (Henderson, 1910, Meyer, 1914). Proprioceptive sensations coming from muscles and proprioceptive sensations of position from muscles, joints and labyrinths affect the respiratory movements. Stimulation of the special senses of sight, hearing and smell and of the sensory endings for cold and heat result in changes in the respiration. This is true also of impulses from the autonomic regions. Centripetal impulses which affect the respiration may come from the whole organism by all pathways.

One must not forget on the other hand the automaticity of the respiratory centers which can maintain up to a certain point alternate inspiration and expiration in the absence of any external influence. This was suggested by Legallois (1812), and recognized later by a large number of investigators (Langendorf, 1888; Schrader, 1887; Lewy, 1891; Lewandovski, 1896; Foa, 1909–11; Winterstein, 1911; Scott and Roberts, 1923; Roberts, 1925; Adrian and Buytendijk, 1931; Finley, 1931; Barcroft, 1934, and many others).

The anatomical localization of the respiratory centers and their functioning have been studied intensively. Flourens supposed that they are located near the tip of the "calamus scriptorius," and Gierke that they are made up of cells connected to the fasciculus solitarius. Mislavski (1885), Aduco (1890) and recently Henderson and Sweet (1929), Finley (1931), Henderson and Craigie (1936), Nicholson (1936), Nicholson and Brezin (1937), Nicholson and Sobin (1938) search the localization of the respiratory centers in the medulla. Worthy of special comment is the paper by Gesell, Bricker and Magge (1936) who studied the electric changes in the function of the medullar respiratory centers. These authors, like Brookhart (1940), state that the neurones which make up the bulbar centers are found more or less scattered through the reticular formation. Pitts, Magoun and Ranson (1939) and Pitts (1940–42) attempt to give a precise anatomical and physiological description of these myelencephalic centers. They distinguish in the cat an "inspiratory area" situated in the ventral reticular formation below the inferior olive, and another dorsal, "expiratory area", somewhat higher, extending beyond and bending over the cephalic portion of the inspiratory area. Excitation of one or the other gives rise respectively to contraction of the muscles of inspiration or of expiration.

There are different physiological categories of centers, the higher centers integrating the function of the lower. This is known since the researches of Markwald (1887–90), Loewy (1888), Langendorf (1888), Luscher (1899), Lewandovsky (1896), etc. Thus Lumsden (1923) and later Stella (1938–39), among others, thought that there is a regulator center in the pons—"pneumotaxic"—which has predominant inhibitory properties and so it is able to control the function of the other subaltern centers. Rijlant (1932), who studied the efferent respiratory electric discharges, established a functional hierarchy among the respiratory centers, possessing different specific properties and located in the medulla and the pons. These centers are not compact nor do they form a strict entity but rather a system of adjacent groups of neurones among which those which have thus far been localized experimentally are of particular functional importance.

In the centers the impulses of various origins are integrated and organized. Motor patterns are formed through the synthesis of the most numerous and various elements and these patterns or functional systems give rise to efferent impulses to the muscles and also to other higher centers both subcortical and cortical, which work in unison and in perfect adjustment with the bulbo-pontile centers.

For a long time it has been supposed that the inspiratory-expiratory alternation was a sequence of reflexes started at the vagal ends of the lungs and related to mechanical stimulus: the state of pulmonary inflation or deflation. The discovery of the Hering-Breuer effect (1868) led to a conception universally accepted. The explanation appears by now incomplete. Expiration, like inspiration, can be active. And side by side with the mechanical stimulation of peripheral receptors should be considered the function of chemo-receptors in different territories, and the intrinsic properties of the respiratory centers at various levels of the neuro-axis, especially the medullo-pontile centers.

Bronk and Fergusson (1935) employ the inscription of the electric "fusillades" by the motor nerves to study this question. They conclude that both inspiration and expiration are active and that the rhythmic impulses of excitations arise automatically in the centers, even in the absence of afferent impulses arriving to them. Gesell (1940), Gesell, Magee and Bricker (1940) and Gesell, Atkinson and Brown (1940) describe the patterns of the electric variations in the respiratory motor nerves; inspiratory impulses present there different patterns from the expiratory ones. These experiments show also that the centers are active in inspiration and expiration and that centrifugal currents may leave such centers independently of any afferent impulses. In the respiratory mechanisms there are centrogenic excitations beside reflexogenic influences. The centers maintain a tone and send out rhythmical impulses over the centrifugal pathways; but they are subject to numerous afferent sensory influences of varied origin and nature, and subject also to their own chemical state and that of the blood passing through them.

Gesell and Hamilton (1941) have made a further study of the reflex mechanism by stimulating three types of nerves: the vagus which they consider proprioceptive, the carotid sinus nerve, chemoceptive, and the saphenous, nociceptive; they

confirm the opinion that the two forces which control the respiration are the activity of the central neurones and the influence exerted on the function of these neurones by afferent impulses arising in receptors of various sorts. The sum total of the influences both central and peripheral which at times interfere and at other times re-enforce each other constitutes the power which drives the central nervous system. An overall summation of these different influences is brought about, an organization which the authors attempt to schematize and by means of which the nervous control of respiration is explained. Gesell and Atkinson (1943) deal with the "motor integration" of the respiration still using the analysis of the respiratory discharges in the diaphragm, in various species of animals. In the centers elementary phenomena are integrated which may possibly be identical or similar along the zoological scale and which become complicated in various ways according to the level of function.

Pitts, Magoun and Ranson (1939) study the functional interrelation of the different respiratory centers in the cat employing the Horsley-Clarke technique. There are in the medulla inspiratory and expiratory bilateral groups of neurones, and fibres of the vagus are connected with them. Pitts (1942) reaches the conclusion that four subsidiary systems are present in the medulla: a, an excitomotor center which sends out periodic discharges; b, a vagal inhibitory system; c, a cerebral inhibitory system, cortico-hypothalamic; d, other excitatory and inhibitory systems. The inhibitory systems, both vagal and cerebral, check the activity of the neurones of the motor system and in this way rhythmical respiration is produced; it is also possible that they vary the rate and depth according to need.

Comore (1944) writes recently: "Probably the most important discovery of the last century has been the realization that respiration is controlled not by stimulation of the medulla alone, neither by reflexes alone, but by proper interaction of both factors. No reflex, no matter how strong, can stimulate respiration if the arterial CO_2 tension has been lowered abnormally (Krogh and Lindhard, 1913; Stella, 1939); no chemical stimulant, no matter how great, can produce rhythmic breathing if the medullary centers have been completely cut off from all nervous influence including that residing in pneumotaxic center (Stella, 1938). Respiratory alterations in general cannot be explained by a single theory but only by a consideration of a number of known and probably many unidentified factors. This has been our thesis, a long time ago, since our initial researches.

We cannot finish without stressing the participation of certain cortical regions in maintaining and adapting the respiratory movements to physiological requirements. Consider for example the adaptation of the respiration to the requirements of spoken language. It is not surprising that speaking as well as silent reading (Bellido, 1922), rhythm (Japelli, 1906; Coleman, 1920), various psychic effects (Allen, 1929–42), emotion and finally consciousness and the will all influence the respiratory movements. Fulton (1943) asserts that the regulation of the respiratory movements is a problem of general physiology in the widest sense. A broad concept of the function will keep us from attributing excessive importance to a few isolated facts and from concluding that extremely simple

and limited mechanisms, such as reflexes divorced from the function of the whole, can determine by themselves the flexible respiratory dynamics. Everything which leads to the recognition of new facts is important. But let us not advance an explanation of the whole, which depends upon innumerable factors, in terms of one elementary function—one receptor, one center and one effector. Investigations which have been carried out without a logical conception of the totality of the neural organization and of the adaptation of the responses have contributed to the obscurity of the conclusions and to a general disorientation. The physiologist must be a careful experimenter, but he must have above all a clear idea of what he is seeking; he must attribute to his observations their exact significance and he must know how to incorporate his results into the general theory.

Therefore when we contemplate in the proper perspective the history of the fundamental discoveries in the physiology of respiration we are surprised at the extreme pains taken by some authors to deny the existence and function of chemo-receptors in the respiratory apparatus. Our purpose in studying this problem was motivated by a leading idea expressed in the thesis of Turró (1914), which stated the functions of a "trophic sense" of chemical nature possessed by many different tissues. After the demonstration of the trophic reflexes of hyperglycemia (Pi-Suñer, 1917) and hyperlipemia (Geelmuyden, 1923; Wertheimer, 1926), it seemed to us that it would be an interesting problem to determine whether a reflex mechanism responsive to peripheral chemical stimuli also acts to control the respiratory movements, which are so easily observed and which offer so much for investigation. The results were positive and the researches of numerous investigators have given a decisively favorable response. From the nose (Allen, 1929; Kerekes, 1935; Deseo and Fodor, 1935) and the larynx and trachea (Lumsden, 1924; Graham, 1939–40) to the tissues themselves," a series of chemo-receptors takes part in the control of the respiratory movements: in the respiratory apparatus at various levels, in the circulatory systems and in other parts of the organism.

The vascular chemo-receptors act effectively in the regulation of the respiratory movements but it cannot be doubted that the location of the chemo-receptors in the respiratory apparatus, at various levels, from the nose to the finest bronchioles and the lungs, is a more strategic one. It is logical to suppose that in mammals, animals which breathe air, receptors in the respiratory apparatus itself have greater functional significance than have vestiges of sensory structures in the vessels, which are useful primarily for respiration in the water (Schmidt, 1938).

Gesell wrote recently: "If ever there was a conviction firmly intrenched in physiology, it was the monopoly of the chemical control of breathing by the respiratory center.... *It proved to be one of physiology's outstanding creeds*.... It was a shaky foundation upon which all of us worked. So when Heymans, *et al.*, produced hyperpnea by a lack of oxygen or an excess of carbon dioxide confined to the aortic (1924–27) and carotid chemo-receptors (1930–32), *he gave us a new outlook on respiration* for which physiology is deeply indebted.... *Both central and peripheral* chemical control were proven to be extremely important."

For over a quarter of a century we have been investigating the functions of the respiratory chemo-receptors. The idea of the rôle of the peripheral chemical stimulation of receptors in the processes of regulation of pulmonary ventilation and respiratory dynamics was enunciated by us in 1918 and at the same time the reflex effects from those stimuli were demonstrated. That year we published our first paper on chemo-receptors in the respiratory apparatus, which control its motor functions, and in 1920 we gave a practical demonstration of the technique employed (the dog with two heads, perfusing the head of the experimental dog with the blood of the donor dog) before the Tenth International Physiological Congress in Paris. Since those first experiments, our research work on this subject has been extensive. Today it is an undisputed truth that *peripheral chemoreceptors* are instrumental in the control and the regulation of the respiratory movements.

REFERENCES

Adrian, E. D. J. Physiol. **61**: 49, 1926.
J. Physiol **79**: 332, 1933.
Adrian, E. D. and F. J. J. Buytendijk. J. Physiol. **71**: 121, 1931.
Aducco, V. Arch. ital. Biol. **13**: 89, 1890.
Alam, M. and F. H. Smirk. J. Physiol. **89**: 372, 1937.
J. Physiol. **90**: 167, 1938.
Allen, W. F. Am. J. Physiol. **88**: 117, 620, 1929.
Am. J. Physiol. **136**: 783. 1942.
Anrep, G. V., W. Pascual and R. Rossler. (Cited in) Anrep and Samaan. *Loc. cit.*, 1932.
Anrep, G. V. and A. Samaan. J. Physiol. **77**: 1, 1933.
Ask-Upmark, E. The carotid sinus and the cerebral circulation. Acta. Psychol. and Neurol. Supl. VI, 1935.
Baglioni, S. Zentralbl. Physiol. **16**: 649, 1903.
Ergebn. Physiol. **11**: 526, 1939.
Bagoury, M. M. and A. Samaan. J. Egypt. med. Assoc. **24**: 211, 1941.
Barcroft, J. F. The respiratory function of the blood. Univ. Press., Cambridge, 1925–28.
Features in the architecture of physiological function. Univ. Press, Cambridge, 1934.
Barcroft, J. F. and R. Margaria. J. Physiol. **72**: 175, 1931.
J. Physiol. **74**: 156, 1932.
Barry, D. T. J. Physiol. **84**: 263, 1935.
Beccari, E. Ann. de Physiol. **10**: 934, 1934.
Arch. intern. Physiol. **39**: 257, 1934.
Ergebn. Physiol. **39**: 257, 1934.
Bellido, J. M. Verbal communication, 1922.
Berns, R. Onderz. Phys. Lab. Utrecht, **3**: 76, 1870.
Bernthal, T. Am. J. Physiol. **109**: 8, 1934.
Am. J. Physiol. **121**: 1, 1938.
Ann. Rev. Physiol. **6**: 155, 1944.
Bernthal, T. and F. J. Schwind. Am. J. Physiol. **143**: 361, 1945.
Bernthal, T. and W. F. Weeks. Am. J. Physiol. **127**: 94, 1939.
Beyne, J., J. Gautrelet and N. Halpern. C. R. Soc. Biol. **113**: 585, 1933.
Binger, C. A. and R. M. Moore. J. Exper. Med. **45**: 633, 643, 655, 1927.
Bogue, J. Y. and H. Rosenberg. J. Physiol. **82**: 353, 1934.
Bogue, J. Y. and G. Stella. J. Physiol. **82**: 23, 1934.
J. Physiol. **83**: 459, 1935.

Bordoni, L. Lo Sperimentale. 1888. Quoted in Luciani. "Fisialojia del l'Uomo. Milano, 1901.
Bouckaert, J. J. and C. Heymans. J. Physiol. **79:** 49, 1933.
Boycott, A. E. and J. S. Haldane. J. Physiol. **37:** 355, 1908.
Boyd, J. D. Contributions to Embriology. Carnegie Inst. **26:** 3, 1937.
Brassfield, C. R. Proc. Soc. exper. Biol. Med. **26:** 833, 1929.
Brewer, M. R. Am. J. Physiol. **120:** 91 1937.
Briscoe, G. J. Physiol. **76:** 52, 1932.
Brodie, T. G. and A. E. Russell. J. Physiol. **26:** 92, 1900.
Bronk, D. W. and L. K. Ferguson. Am. J. Physiol. **110:** 708, 1935.
Bronk, D. W., L. K. Ferguson, R. Margaria and D. Y. Solandt. Am. J. Physiol. **117:** 237, 1936.
Bronk, D. W. and G. Stella. J. Cell. comp. Physiol. **1:** 113, 1932.
Brookhart, J. M. Am. J. Physiol. **129:** 709, 1940.
Brookhart, J. M. and E. H. Steffensen. Am. J. Physiol. **115:** 357, 1936.
Bulbring, E. and J. Whitteridge. J. Physiol. **102:** 23, 1943–44.
Cachovsky, P. Doctoral thesis, Petrograd, 1899.
Campbell, J. M. H., C. G. Douglas, J. S. Haldane and F. G. Hobson. J. Physiol. **46:** 301, 1913.
Campbell, J. M. H., C. G. Douglas and F. G. Hobson. J. Physiol. **48:** 303, 1914.
Camus, L., H. Benard and F. P. Merklen. Compt. Rend. Soc. Biol. **115:** 614, 1626, 1934.
Cannon, W. B., J. T. Lewis and S. W. Britton. Am. J. Physiol. **77:** 326, 1926.
Castro, F. de. Trav. Lab. Recher. Biol. Univ. de Madrid **24:** 365, 1926.
Trav. Lab. Recher. Biol. Univ. de Madrid **25:** 331, 1927–28.
Cheshcov, A. M. Doctoral Thesis, Petrograd, 1902.
Christensend, E. H., A. Krogh and J. Lindhard. Skand. Arch. Physiol. **74:** Sup. 10 1, 1936.
Christie, R. V. Quart. J. Med. **7:** 421, 1938.
Churchill, E. D. and O. Cope. J. exper. Med. **49:** 531, 1929.
Coleman, W. M. J. Physiol. **54:** 213, 1920–21.
Comroe, J. H., Jr. Am. J. Physiol. **127:** 176, 1939.
Am. J. Physiol. **139:** 490, 1943.
Physiol. Rev. **24:** 319, 1944.
Comroe, J. H., Jr. and C. F. Schmidt. Am. J. Physiol. **121** : 75, 1938.
Am. J. Physiol. **138:** 536, 1943.
Cordier, D. and C. Heymans. "Le Centre Respiratoire" Assoc. Physiol., 1935.
Ann. Phys. et Phys.-Chem. Biol. **11:** 535, 1935.
Cromer, S. P. and A. C. Ivy. Proc. Soc. Exper. Biol. and Med. **28:** 565, 1931.
Danielopolu, D. Ztschr. exper. Med. **63:** 130, 1928.
Danielopolu, D., A. Asland, I. Marcou, G. Proca and E. Manescu. Pres. Med. 1585, 1927.
Danielopolu, D., E. Manescu and G. Proca. Ztschr. exper. Med. **63:** 139, 143, 1928.
Danielopolu, D., A. Asland and I. Marcou. J. Physiol. et Path. gen. **31:** 338, 1933.
Dautrebande, L. Les Echanges Respiratories, Paris, 1930.
Volume Jubilaire en honn. du Prof. Demoor, 1937.
Davies, R., J. S. Haldane and E. L. Kennaway. J. Physiol. **54:** 32, 1920.
Decharneux, G. Compt. Rend. Soc. Biol. **116:** 352, 1934.
Desseo, D. von and L. Fodor. Pflüger's Arch. **236:** 554, 1935.
Dill, D. B. Life, heat and altitude, 1938.
Dirkin, M. N. J. and H. A. E. van Dishoeck. Pflüger's Arch. **238:** 713, 1937.
Donders, F. C. Ztschr. f. ration. Med. **3:** 287, 1853.
Douglas, C. G. and R. H. Harvard. J. Physiol. **74:** 471, 1932.
Dripps, R. D. and J. H. Comroe, Jr. Am. J. Med. Sci. **20:** 681, 1944.
Du Bois-Reymond, R. and R. Katzenstein. Arch. f. Physiol. **513:** 436, 1901.

Dumke, P. R., C. F. Schmidt and H. P. Chiodi. Am. J. Physiol. 133: 1, 1941.
Dunn, J. S. Quart. J. Med. 13: 129, 1920.
Eppinger, H., L. von Papp and H. Schwartz. Ueber das Asthma Cardialis. Berlin, 1924.
v. Euler, U. S. and G. Liljestrand. Skand. Arch. Physiol. 74: 101, 1936; 77: 101, 1937.
v. Euler, U. S. Skand. Arch. Physiol. 80: 94, 1938.
v. Euler, U. S., G. Liljestrand and Y. Zottermann. Skand. Arch. Physiol. 83: 132, 1939. Acta. Physiol. Skand. 2: 1, 1941.
Finley, K. H. Arch. Neurol. and Psych. 26: 754, 1931.
Fleisch, A. Ztschr. allg. Physiol. 19: 269, 1921.
Arch. Exper. Path. Pharmac. 105: 17, 1922.
Pflüger's Arch. 219: 706, 1928.
Pflüger's Arch. 223: 509; 224: 390; 225: 26, 1930.
Ergebn. Physiol. 36: 249, 1934.
XVI Congr. Intern. Physiol. Zurich, 1938.
Foa, C. Arch. Fisiol. 7: 195, 1909.
Arch. Fisiol. 8: 101, 1910.
Arch. Fisiol. 9: 453, 1911.
Fox, H. M. and M. L. Johnson. J. exper. Biol. 11: 1, 1934.
Francois-Franck. Arch. de Physiol. 2: 546, 1891.
Frankel and Geppert. C. R. Acad. Science 96: 1740, 1883.
Fredericq, L. Trav. du Lab. de Physiol. de Liege 3: 1, 1890.
Frey, W. Klin. Wchnschr. 672, 1923.
Frey, W. and J. Hagemann. Ztschr. ges. exper. Med. 25: 271, 1921.
Fulton, J. F. Physiology of the nervous system. 2d ed. Oxford. Univ. Press, 1943.
Gad, J. Arch. f. Physiol. 1: 1880.
Garcia Banus, M., H. H. Corman, V. P. Perlo and G. L. Popkin. Am. J. Physiol. 142: 121, 1944.
Gayet, R., D. Bennati and D. Quivy. Arch. intern. de Pharm. et Therap. 50: 129, 1935.
Gaylor, J. B. Brain 57: 143, 1934.
Gellhorn, E. and E. H. Lambert. Illinois Med. and Dental Mon., 1939.
Geelmuyden, H. C. Ergebn. Physiol. 21: 274, 1923.
Gemmill, C. L., E. M. K. Geiling and D. L. Reeves. Am. J. Physiol. 109: 709, 1934.
Gemmill, C. L. and D. L. Reeves. Am. J. Physiol. 105: 487, 1933.
Geppert, J. Ztschr. J. Klin. Med. 15: 208, 307, 1888–9.
Geppert. J. and N. Zuntz. Pflüger's Arch. 42: 189, 1888.
Gesell, R. Proc. Soc. exper. Biol. and Med. 20: 345, 1922–23.
Physiol. Rev. 5: 551, 1925.
Am. J. Physiol. 115: 168, 1936.
Am. J. Physiol. 116: 228, 1936.
Annual Rev. Physiol. 1: 185, 1939.
Science. 91: 229, 1940.
Gesell, R. and A. K. Atkinson. Am. J. Physiol. 139: 745, 1943.
Gesell, R., A. K. Atkinson and R. C. Brown. Am. J. Physiol. 128: 629, 1940.
Gesell, R., J. Bricker and C. Magee. Am. J. Physiol. 117: 423, 1936.
Gesell, R., H. Krueger, G. Gorham and T. Bernthal. Am. J. Physiol. 94: 300, 1930.
Gesell, R. and M. A. Hamilton. Am. J. Physiol. 133: 694, 1941.
Gesell, R. and J. Lapides. XVI. Intern. Congr. Phys. Zurich, 1938.
Gesell, R., J. Lapides and M. Levin. Am. J. Physiol. 130: 155, 1940.
Gesell, R., C. S. Magee and J. W. Bricker. Am. J. Physiol. 128: 615, 1940.
Gesell, R. and C. Moyer. Proc. Soc. exper. Biol. and Med. 31: 1089, 1934.
Quart. J. Exper. Physiol. 24: 332, 1934–35.
Quart. J. Exper. Physiol. 25: 13, 1935.
Am. J. Physiol. 119: 55, 1937.

GESELL, R., C. A. MOYER AND J. B. KITTRICK. Am. J. Physiol. **136:** 486, 1942.
GIBBS, F. A., E. L. GIBBS, W. G. LENNOX AND L. F. NIMS. J. Aviat. Med. **14:** 520, 1943.
GITHENS, T. S. AND S. J. MELTZER. Proc. Soc. Exper. Biol. and Med. **12:** 64, 1916.
GLASSNER, K. Wien. Klin. Wchnschr. **920:** 1906.
GOLLWITZER-MEYER, K. Biochem. Ztschr. **151:** 54, 1924.
Pflüger's Arch. **234:** 342, 1934.
GOLLWITZER-MEYER, K. AND R. SCHULTE. Pflüger's Arch. **229:** 251, 1931.
GOORMAGHTIGH, H. J. Med. de Bruxelles, 1935.
GRAHAM, J. D. P. J. Physiol. **97:** 525, 1939–40.
GREEN, M. F. AND A. DE GROAT. Am. J. Physiol. **112:** 488, 1935.
GREENE, J. A. AND L. W. SWANSON. Arch. int. Med. **61:** 720, 1938.
HAGGARD, H. W. AND Y. HENDERSON. Am. J. Physiol. **61:** 289, 1922.
HALDANE, J. S. Respiration. Yale Univ. Press, 427. pp., 1922.
HALDANE, J. S., A. M. KELLAS AND E. L. KENNAWAY. J. Physiol. **53:** 181, 1919.
HALDANE, J. S. AND E. P. POULTON. J. Physiol. **37:** 390, 1909.
HALDANE, J. S. AND J. G. PRIESTLEY. J. Physiol. **32:** 225, 1905.
Respiration. Oxford Clarendon Press, 2nd. ed. **1935.**
HAMMOUDA, M., A. SAMAAN AND W. H. WILSON. J. Physiol. **101:** 446, 1942–43.
HARRISON, T. R. Failure of the circulation. Baltimore, 1939.
HARRISON, W. G. JR., J. A. CALHOURN AND T. R. HARRISON. Am. J. Physiol. **100:** 68, 1932.
HARRISON, T. R., W. G. HARRISON JR., J. A. CALHOURN AND J. P. MARSH. Arch. int Med' **50:** 690, 1932.
HASSELBACH, K. A. Biochem. Ztschr. **46:** 403, 1912.
HASSELBACH, K. A. AND J. LINDHARD. Biochem. Ztschr. **68:** 265, 1915.
Biochem. Ztschr. **74:** 48, 1916.
HEAD, H. J. Physiol. **10:** 1, 1889.
Pflüger's Arch. **130:** 337, 1903.
HEERDT, P. F. AND B. J. KRIJGSMAN. Ztschr. verg. Phys. **27:** 29, 1939.
HEGER, P. Beitr. Phys. Ludwig gewidmet, 193, 1887.
HENDERSON, L. J. The Oxford Medicine **1:** 471, 1919.
Blood. Yale University Press, New Haven 1928.
HENDERSON, Y. Am. J. Physiol. **21:** 126, 1908.
Am. J. Physiol. **24:** 66, 1909.
Am. J. Physiol. **25:** 310, 385, 1910.
Am. J. Physiol. **27:** 152, 1910.
Am. J. Physiol. **46:** 533, 1918.
Physiol. Rev. **4:** 329, 1924.
Physiol. Rev. **5:** 31, 1925.
Adventures in respiration Baltimore, 1938.
HENDERSON, Y. AND H. E. CRAIGIE. Am. J. Physiol. **115:** 520, 1936.
HENDERSON, Y. AND L. A. GREENBERG. Am. J. Physiol. **109:** 51, 1934.
HENDERSON, Y. AND H. W. HAGGARD. J. Biol. Chem. **33:** 333, 345, 355, 365, 1918.
HENDERSON, V. W. AND SWEET. Am. J. Physiol. **91:** 96, 1929.
HERING, E. Sitzungsb. Akad. Wissen. Wien **64:** Teil 2, 223, 1871.
HERING, H. E. Pflüger's Arch. **206:** 721, 1924.
Ztschr. Kreislauffors. **19:** 410, 1927.
Die Karotissinus Reflexe auf Herz und Geffasse. Dresden, 1927.
Der Blutzuglertonus in seiner Bedeutung fur den Sympathicus und Parasympathicus Tonus. Leipzig, 1932.
HERING, E. AND R. BREUER. Abhand. Wien Akad. Naturwiss. **57:** 672, 1868.
HERMANN, L. AND T. ESCHER. Pflüger's Arch. **3:** 3, 1870.
HERMANN, H., F. JOURDAN AND J. VIAL. J. Physiol. et Path. gen. **32:** 343, 1934.

HESS, W. R. Pflüger's Arch. **168**: 191, 1917.
Pflüger's Arch. **213**: 163, 1926.
Ztschr. Physiol. Chem. **117**: 284, 1921.
Ergebn. inn. Med. **23**: 1, 1923.
Pflüger's Arch. **226**: 198, 1930.
Die Regulierung der Atmung. Leipzig, 1931.
HESS, W. R. AND O. A. M. WYSS. Pflüger's Arch. **237**: 761, 1936.
HEYMANS, C. Verhand. deutsch. Gess. Kreislaufforschung **1**: 92, 1928.
Arch. int. Pharmac. Therap. **35**: 296, 1929.
Le Sinus Carotidien et les autres zones vasosensibles reflexogenes, Louwain, 1929.
Ergebn. der Physiol. **28**: 244, 1929.
XII Intern. Congr. Physiol. Boston, 1929.
Am. J. Physiol. **90**: 387, 1929.
Rev. Med. de Barcelona **14**: 417, 1930.
Reun. Plen. de la Soc. de Biol. Paris, 1931.
HEYMANS, C. AND J. J. BOUCKAERT. C. R. Soc. Biol. **100**: 199, 1929.
J. Physiol. **69**: 254, 1930.
C. R. Soc. Biol. **103**: 31, 498, 1932.
Association des Physiologistes. Lille, 1932.
Ann. Physiol. **8**: 330, 1932.
Ergebn. Physiol. **41**: 28, 1939.
HEYMANS, C., J. J. BOUCKAERT AND L. DAUTREBANDE. C. R. Soc. Biol. **105**: 881, 1930.
C. R. Soc. Biol. **106**: 52, 54, 469, 1276, 1279, 1931.
Arch. Intern. Pharm. Therap. **39**: 400, 1930.
Arch. Intern. Pharm. Therap. **40**: 54, 1931.
J. Physiol. **71**: 5. P. 1931.
Pflüger's Arch. **230**: 283, 1932.
HEYMANS, C., J. J. BOUCKAERT, U. S. EULER AND L. DAUTREBANDE. C. R. Soc. Biol. **109**: 566, 1932.
HEYMANS, C., J. J. BOUCKAERT AND R. REGNIERS. Le Sinus Carotidien et la zono homologue Cardio-aortique. Paris, 1933.
HEYMANS, C., J. J. BOUCKAERT AND A. SAMAAN. C. R. Soc. Biol. **118**: 1246, 1935.
HEYMANS, J. F. AND C. HEYMANS. Arch. Intern. Pharm. Therap. **32**: 9, 1926.
Arch. Intern. Pharm. Therap. **33**: 273, 1927.
C. R. Soc. Biol. **99**: 633, 1928.
HEYMANS, C. AND A. LADON. Arch. Intern. Pharm. et Therap. **30**: 415, 1925.
HEYMANS, C. AND P. RIJLANT. C. R. Soc. Biol. **113**: 69, 1933.
HILL, L. AND M. FLACK. J. Physiol. **37**: 77, 1908.
HOFMANN, P. Deutsch. Klin. **4**: 2, 17, 1900.
HOFMANN, P. AND C. R. KELLER. Ber. der Physiol. **221**: 296, 1929.
HOLLINSHEAD, W. H. AND C. H. SAWYER. Am. J. Physiol. **144**: 79, 1945.
HOOKER, D. R., D. W. WILSON AND H. CONNET. Am. J. Physiol. **43**: 351, 1917.
HOUSSAY, B. A. AND O. ORIAS. Compt. Rend. Soc. Biol. **116**: 896, 1934.
HUDOVERNIG, L. J. Psych. und Neurol. **9**: 234, 1907.
IRVING, L. Am. J. Physiol. **122**: 207, 1938.
Physiol. Rev. **19**: 112, 1939.
IRVING, L. AND M. S. WELCH. Quart. J. Exper. Physiol. **25**: 121, 1935.
JAPELLI, L. Arch. di Fisiol. **3**: 170, 1906.
JOHANSSON, J. E. Skand. Arch. Physiol. **5**: 20, 1895.
JONGBLOED, M. J. Ann. Physiol. **12**: 457, 1936.
JOSEPH, D. R. Am. J. Physiol. **59**: 491, 1922.
KEREKES, G. Acta Oto-Laryng. **21**: 438, 1935.
KROGH, A. The comparative physiology of respiratory mechanism. Univ. Penna. Press, Philadelphia, 1941.

KROGH, A. AND J. LINDHARD. J. Physiol. **47**: 112, 1913.
J. Physiol. **51**: 182, 1917.
J. Physiol. **53**: 431, 1920.
LANGENDORF, O. Arch. Anat. Physiol. **8**: 285, 1887.
Arch. Anat. Physiol. **9**: 303, 1888.
LARSELL, O. J. Comp. Neurol. **33**: 105, 1921.
Textbook of neuro-anatomy and the sense organs. New York, 1939.
LARSELL, O. AND E. G. BURGET. Am. J. Phys. **70**: 311, 1924.
LARSELL, O. AND R. MASSON. J. Comp. Neurol. **33**: 509, 1921.
LATSCHENBERGER, J. AND A. DEAHNA. Pflüger's Arch. **12**: 157, 1876.
LEGALLOIS, C. F. Experiences sur le Principe de la Vie. Paris, 1812.
LEWANDOWSKY, M. A. Arch. Anat. Physiol. **17**: 195, 483, 1896.
LENNOX, W. G. AND E. L. GIBBS. J. Clin. Investigation **11**: 1155, 1932
LEWY, A. Pflüger's Arch. **49**: 406, 1891.
LIM, B. K. J. Physiol. **14**: 467, 1893.
LINDORTH, A. Arch. Zool. **30-B**: 3, 1, 1938.
LOEWY, A. Pflüger's Arch. **42**: 245, 273, 1888.
LUCIANI, L. Fisiologia de l'Uomo. Milano, 1901.
LUCIANI, L. Lo Sperimentale, 1888. Cita de Luciani. Loc. cit.
Lo Sperimentale, 1888. Cita de Luciani. Loc. cit.
LUMSDEN, T. J. Physiol. **57**: 153, 354, 1923.
J. Physiol. **58**: 81, 111, 1924.
LUSCHER, F. Ztschr. Biol. **38**: 499, 1899.
MAGOUN, H. W. Am. J. Physiol. **122**: 530, 1938.
McGINTY, D. A. AND R. GESELL. Am. J. Physiol. **75**: 70, 1925.
MARKWALD, M. Ztschr. Biol. **23**: 149, 1887.
Ztschr. Biol. **26**: 259, 1890.
MARKWALD, M. Ztschr. Biol. **26**: 259, 1890.
MARSHALL, E. K. JR. AND M. ROSENFELD. J. Pharmacol. and Exper. Therap. **59**: 222, 1937.
MAYER, A., H. MAGNE AND L. PLANTEFOL. Comp. Rend. Acad. Science **170**: 1347, 1920.
MEEK, W. J. Am. J. Physiol. **68**: 309, 1923–24.
MEYER, A. L. J. Physiol. **48**: 47, 1914.
MEYER, H. Ztschr. verg. Physiol. **22**: 435, 1935.
MIES, H. Ztscht. ges. Exper. Med. **85**: 282, 1932.
MISLAVSKI, N. A. Arch. Psychiat. **6**: 138, 1885.
MOISEJEFF, E. Ztschr. ges. Exper. Med. **53**: 696, 1927.
MONIZ DE BETTENCOURTA, J., M. RODRIGUES CARDOSO AND F. PAES DE VASCONCELLOS.
C. R. Soc. de Biol. **129**: 701, 1938.
MOSSO, A. Arch. ital. Biol. **7**: 52, 1885.
MOYER, C. A. AND H. K. BEECHER. Am. J. Phys. **136**: 13, 1942.
MULINOS, M. G. J. Pharmacol. and Exper. Therap. **51**: 135, 1934.
NIELSEN, M. Skand. Arch. Physiol. **74**: Supl. **10**: 87, 1936.
NICHOLSON, H. C. Am. J. Physiol. **115**: 402, 1936.
NICHOLSON, H. C. AND D. BREZIN. Am. J. Physiol. **118**: 441, 1937.
NICHOLSON, H. C. AND S. SOBIN. Am. J. Physiol. **123**: 766, 1938.
NONIDEZ, J. F. Am. J. Anat. **57**: 259, 1935.
J. Anat. **70**: 215, 1936.
OLTHOF, H. J. Ztschr. Verg. Physiol. **21**: 534, 1936.
OWEN, H. AND R. GESELL. Proc. Soc. Exper. Biol. and Med. **28**: 765, 1931.
PAGANO, L. Arch. ital. Biol. **33**: 1, 1900.
PATRIDGE, R. C. J. Cellular Comp. Physiol. **2**: 367, 1933.
PATTERSON, W. D. J. Physiol. **66**: 323, 1928.
PAVLOV, I. P. Proc. Russian Med. Soc. 1895.
Proc. Russian Med. Soc., 1896.

PELECOVICH, M. Am. J. Physiol. **99**: 357, 1932.
PENITSCHKA, W. Ztschr. Mikro-Anat. Forsch. **24**: 1931.
PERMAN, E. Ztschr. Anat. **71**: 382, 1924.
PETERS, F. Ztschr. Verg. Physiol. **25**: 591, 1938.
PFLÜGER, E. Pflüger's Arch. **1**: 61, 1868.
PI-SUÑER, A. Treb. Soc. Biol. de Barcelona **3**: 54, 1918.
Treb. Soc. Biol. de Barcelona **4**: 76, 253, 1919.
Treb. Soc. Biol. de Barcelona **5**: 71, 1920.
Treb. Soc. Biol. de Barcelona **7**: 41, 1922.
Treb. Soc. Biol. de Barcelona **10**: 120, 1925.
Treb. Soc. Biol. de Barcelona **12**: 161, 1928.
Treb. Soc. Biol. de Barcelona **14**: 400, 1931.
Treb. Soc. Biol. de Barcelona **17**: 281, 1934.
Journ. Phys. Path. Gen. **18**: 702, 1919.
C. R. Soc. Biol. **90**: 813, 821, 1924.
XII Intern. Congr. Phys. Boston, 1929.
Am. J. Physiol. **90**: 473, 1929.
Reunion pleniere Soc. Biol. Paris, 1931.
C. R. Soc. Biol. **107**: 1324, 1931.
XVI Intern. Congr. Physiol. Zurich, 1938.
Anal. Inst. Med. Exper. Caracas **1**: 140, 1942.
PI-SUÑER, A. AND J. M. BELLIDO. Treb. de la Soc. de Biol. **4**: 311, 1910.
X Intern. Congr. Physiol. Paris, 1920.
J. Phys. et Path. Gen. **19**: 214, 1921.
PI-SUÑER, A., F. DOMENECH ALASINA AND B. BENAIGES. Treb. de la Soc. de Biol. **15**: 125 1933.
C. R. Soc. Biol. **114**: 1419, 1933.
PI-SUÑER, A. AND J. PUCHE. Treb. Soc. Biol. de Barcelona **7**: 249, 1922.
Treb. Soc. Biol. de Barcelona **8**: 137, 143, 1923.
Treb. Soc. Biol. de Barcelona **11**: 313, 1927.
Treb. Soc. Biol. de Barcelona **12**: 116, 1928.
Comp. Rend. Soc. Biol. **103**: 735, 1930.
PI-SUÑER, A., J. PUCHE, AND A. RAVENTOS. Treb. Soc. Biol. **13**: 169, 1930.
PI-SUÑER, A. AND A. RAVENTOS. Treb. Soc. Biol. de Barcelona **14**: 476, 1931.
Treb. Soc. Biol. de Barcelona **15**: 43, 1932.
Treb. Soc. Biol. de Barcelona **16**: 64, 1933.
PI-SUÑER, J. Trab. Inst. Fisiol. Barcelona **3**: 122, 1929.
PIKE, R. H. AND L. M. COMBS. Am. J. Physiol. **59**: 472, 1922.
Science **56**: 671, 691, 1922.
PIRAS, A. Arch. fisiol. **20**: 359, 1922.
PITTS, R. F. J. Comp. Neurol. **72**: 605, 1940.
Am. J. Physiol. **134**: 192, 1941.
J. Neurophysiol. **5**: 75, 403, 1942.
PITTS, R. F., H. W. MAGOUN AND S. W. RANSON. Am. J. Physiol **126**: 673, 289, 1930.
Am. J. Physiol. **127**: 654, 1939.
PORTER, W. T. AND L. H. NEWBURGH. Am. J. Physiol. **42**: 175, 1916–17.
POWERS, E. B. AND R. T. CLARK, JR. Am. J. Physiol. **138**: 104, 1942.
PUCHE, J. C. R. Soc. Biol. **88**: 617, 1923.
QUINCKE, H. I. Berlin. klin. Wchnschr. 189, 1875.
REIN, H. XVI Intern. Congr. Physiol. Zurich, 1938.
RIJLANT, P. Arch. int. Physiol. **35**: 326, 1932.
ROBERTS, F. J. Physiol. **59**: 460, 1925.
ROSENTHAL, J. In Handbuch der Physiology by Hermann, Vol. **4**: Teil 2, 1880.
SAMAAN, A. AND G. STELLA. J. Physiol. **85**: 309, 1935.

SCHMIDT, C. F. Am. J. Physiol. **84**: 202, 1928.
Am. J. Physiol. **102**: 94, 119, 1932.
In Macleod's Physiology in modern medicine, 8 ed., 1938.
J. Lab. Clin. Med. **26**: 223, 1940; 9 ed., 1941.
SCHMIDT, C. F. AND W. B. BENSON HARER. J. Exper. Med. **37**: 69, 1923.
SCHMIDT, C. F. AND J. H. COMROE, JR. Physiol. Rev. **20**: 115, 1940.
Ann. Rev. Physiol. **3**: 151, 1941.
SCHMIDT, C. F., J. H. COMROE, JR. AND R. D. DRIPPS, JR. Proc. Soc. Exper. Biol. and Mc'. **42**: 31, 1939.
SCHMIDT, C. F., P. R. DUMKE AND R. D. DRIPPS, JR. Am. J. Physiol. **128**: 1939–40.
SCHNEIDER, E. C., D. TRUESDELL AND R. W. CLARKE. Am. J. Physiol. **78**: 393, 1926.
SCHRADER, M. E. G. Pflüger's Arch. **41**: 75, 1887.
SCHUMACHER, S. VON. Ann. Anat. Path. **21**: 1, 1902.
SCHWIEGK, H. Pflüger's Arch. **236**: 206, 1935.
SCOTT, F. H. J. Physiol. **37**: 301, 1908.
SCOTT, F. H., C. C. GAULT AND R. KENNEDY. Am. J. Physiol. **59**: 471, 1922.
SCOTT, J. M. D. AND F. ROBERTS. J. Phys. **58**: 168, 1923.
SELLADURAI, S. AND S. WRIGHT. Quart. J. Exper. Physiol. **22**: 285, 333, 1932.
SHARPEY-SCHAFFER, E. J. Physiol. **75**: 130, 1932.
SICILIANO, R. Arch. ital. Biol. **33**: 338, 1900.
SIDNEY HARRIS, A. Am. J. Physiol. **143**: 140, 1945.
SMYTH, D. H. J. Physiol. **87**: 70, 1936.
J. Physiol. **88**: 425, 1937.
SOLLMANN, T. AND E. D. BROWN. Am. J. Physiol. **30**: 88, 1912.
SOMER, E. DE AND J. F. HEYMANS. J. Physiol. et Path. Gen. **14**: 1139, 1912.
STADIE, W. A., J. H. AUSTIN AND H. W. ROBINSON. J. Biol. Chem. **66**: 901, 1925.
SPALATO, F. AND R. CONSIGLIO. I Nervi Vasi-Sensitivi, 1886.
STARLING, E. H. Principles of human physiology, 2 ed. London, 1915.
STELLA, G. Quart. J. Exper. Physiol. **25**: 145, 1935.
J. Physiol. **93**: 263, 1938.
J. Physiol. **95**: 365, 1939.
STEWART, G. N. Am. J. Physiol. **20**: 407, 1907.
SUNDER-PLASSMANN, P. Ztschr. Anat. Entwick. **93**: 567, 1930.
Deutsch. Ztschr. Chir. **240**: 1933.
TAYLOR, H. J. Physiol. **69**: 124, 1930.
TELLO, F. Trav. Lab. Rech. Biol. Univ. de Madrid **22**: 295, 1924.
TOURNADE, A. C. R. Soc. Biol. **109**: 1123, 1932.
TOURNADE, A. AND MALMEJAC. C. R. Soc. Biol. **106**: 444, 1931.
TOURNADE, A. AND ROCCHISANI. C. R. Soc. Biol. **115**: 1639, 1934.
TRAUBE, E. L. Gess. Beitr. Physiol. **1**: 1871.
TSCHERMAK, A. Jenae Ztschr. Med. und Naturwiss. **2**: 384, 1866.
TURRO, R. Les Origines de la Connaissance, Paris, 1914.
VARALDI, S. Arch. ital. di Biol. **19**: 243, 1893.
VERWORN, M. Arch. Anat. u. Physiol. **65**: 65, 1903.
VIERORDT, O. In Wagner's Handworterbuch del Physiologie, 1844.
VOLKMANN, R. Muller's Arch. f. Anat. und Physiol. **21**: 342, 1841.
WAELE, H. AND J. VAN DE VELDE. Arch. int. Physiol. **50**: 33, 1940.
WATT, J. G., P. R. DUMKE AND J. H. COMROE, JR. Am. J. Physiol. **138**: 610, 1943.
WENKEBACH, E. AND WINTERBERG. Die Unregelmassige Herztatigkeit. Lepzig, 1927.
WERTHEIMER, E. Pflüger's Arch. **213**: 262, 1926.
WESTERLUND, A. Skand. Arch. Physiol. **18**: 263, 1906.
WINDER, C. V. Am. J. Physiol. **106**: 28, 1933.
Am. J. Physiol. **118**: 389, 1937.
Am. J. Physiol. **136**: 200, 1942.

WINDER, C. V., T. BERNTHAL AND W. F. WEEKS. Am. J. Physiol. **124**: 238, 1938.
WINDER, C. V. AND H. O. WINDER. Am. J. Phys. **105**: 337, 1933.
WINDER, C. V., H. O. WINDER AND R. GESELL. Am. J. Physiol. **105**: 311, 1933.
WINTERSTEIN, H. Pflüger's Arch. **138**: 159, 1911.
Pfluger's Arch. **187**: 293, 1921.
Ztschr. vergl. Physiol. **2**: 315, 1925.
WINTERSTEIN, H. AND FRUHLING. Pflüger's Arch. **234**: 187: 1934.
WITT, D B., L. N. KATZ AND L. KOHN. Am. J. Physiol. **107**: 213, 1934.
WRIGHT, S. Quart. J. Exper. Physiol. **24**: 169, 1934,
Quart. J. Exper. Physiol. **26**: 63, 1936.
Quart. J. Exper. Physiol. **28**: 33, 1938.
ZOTTERMANN, Y. Skand. Arch. Physiol. **72**: 73, 1935.

THE ACTION OF INSULIN

J. P. BOUCKAERT AND CHR. DE DUVE

Physiological Laboratory, University of Louvain, Belgium

In spite of the enormous amount of both experimental and critical work devoted in earlier years to the problem of insulin action, and in spite of the apparently convincing demonstration furnished by earlier work that insulin facilitates both glucose oxidation and storage as muscle glycogen, too many facts are still disputed, too many conclusions remain questioned, to allow a satisfactory agreement to be reached on the subject. Since then, a great step forward has been made, thanks particularly to the results achieved by Soskin and other American authors; new light has been thrown simultaneously on the biochemical aspect of the problem by the Coris' more recent work on glucose and glycogen phosphocatalytic metabolism. To these important results we now wish to add our own contribution to the subject.

It may be recalled that the problem of insulin action has been studied uninterruptedly since as early a date as 1923, in the physiological laboratory at Louvain University. The earlier experiments by Bouckaert, De Nayer and their co-workers already pointed against the current views on the subject; as to our later results, not only do they conclusively disprove the classical muscular theory of insulin action, but, jointly with Soskin's and other recent data, they afford sufficient evidence for reconsidering the whole bulk of experimental work on which it is based. Out of the new facts thus brought to light, a different theory arises, which provides a more comprehensive and satisfactory picture of insulin action than has yet been put forward. As has been shown in the detailed review on carbohydrate metabolism, diabetes and insulin, recently published by one of us (27), practically all the results obtained until now on this subject may be gathered, when properly interpreted, in support of this theory.

Our essential aim, in this paper, is to summarize our most important results and to point out briefly how they affect the general theory of insulin action. For a more detailed account we wish to refer to the work of de Duve (27), already mentioned, where the more recent literature has been extensively discussed.

EXPERIMENTAL RESULTS AND DISCUSSION. a. *The importance of the blood sugar level.* Some experiments by Wierzuchowski, Soskin, Lundsgaard, amongst others, have served to emphasize a relationship between blood-sugar level and blood-sugar consumption. Such a relationship has been shown to exist by De Nayer (17) between the amount of glucose deposited as muscle glycogen under insulin action and the blood-sugar level at which the insulinized animal is maintained by continuous intravenous infusion of glucose; we ourselves have observed the quantity of glucose disappearing under the influence of insulin to be directly proportional to the mean glycaemia of both rabbits and dogs receiving insulin and glucose intravenously (19).

The mathematical function relating blood-sugar consumption to blood-sugar level has been the object of some disagreement. According to Lundsgaard and

co-workers (44) the function is linear in a range of blood sugar from 0.20 to 0.60 per cent; Wierzuchowski (64), Soskin and Levine (60) both find the curve to be parabolic. In the latter work, however, the total carbohydrate utilization is considered and not the blood-sugar disappearance only. On the other hand, the data obtained by Wierzuchowski can be satisfactorily put on a straight line in the blood-sugar range in which Lundsgaard finds the relationship to be linear. Our own results show that the linear relationship still appears at lower blood-sugar levels in the case of the insulinized animal. From this fact it may be assumed that the relationship really is linear, but that the curve tends to become parabolic at low or very high blood-sugar levels. This would be due, in the latter case, to saturation of the enzymes responsible for glucose disappearance, in the former to the fact that hepatic glycogenesis[1] replaces to some extent infused glucose in compensating peripheral consumption, and that the exogenous supply, by which consumption is measured, must be accordingly restricted.

The relationship between *glucose administered* and blood-sugar level would then become linear from the moment when hepatic glycogenesis became completely inhibited, in other words from the hepatic threshold[2] upwards, to the point where the curve is again flattened by saturation of the enzymes. This is indeed the case in Wierzuchowski's experiment, if we assume that, in accordance with Soskin, the normal value for the hepatic threshold lies between 0.15 and 0.20 per cent. Our own results would then support the view that insulin considerably lowers the hepatic threshold, allowing the real relationship to be observed at much lower rates of glycaemia, where the curve of the non-insulinized animal already assumes a parabolic aspect. As will be shown, this view is conclusively demonstrated by our further experiments.

To our knowledge, no explanation has yet been put forward to account for the observed relationship between blood-sugar consumption and blood-sugar level. We ourselves consider it as a *necessary consequence of the law of mass-action.* Whatever the tissues concerned and whatever the final product of this transformation may be, glucose consumption must necessarily be equivalent to glucose transformation and must therefore be due to chemical reactions. The *initial reaction*, by which glucose enters the metabolic processes will occur at the expense of glucose itself; *its velocity will be increased by an increase in the concentration of glucose.* Secondly, the subsequent reactions will also be accelerated and a total increase of the metabolic turnover of glucose will occur. In accordance with this

[1] To prevent confusion, we wish to point out that, throughout this paper, we have used the expression "hepatic glycogenesis" in the sense of "sugar formation by the liver," whatever the origin of the sugar formed. Hepatic glycogenesis will therefore cover both "hepatic glycogenolysis" (sugar formation from glycogen) and "glyconeogenesis" (sugar formation from non carbohydrate sources such as protein and, possibly, fat). Synthesis of glycogen will be referred to as "glycogen formation." It should be further noted that the word "glycogenolysis" really means "glycogen breakdown." When applied to the liver, it usually means "glycogen breakdown to glucose;" but many authors also refer to "muscular glycogenolysis" in relation with glycogen breakdown to lactic acid. In the latter case, we have adopted the expression "glycogen breakdown."

[2] The hepatic threshold is the value of the blood-sugar level, at which sugar formation becomes completely inhibited in the liver and gives place to sugar retention.

line of reasoning, the mathematical function relating blood-sugar disappearance to blood-sugar level should have a specific significance, a linear curve indicating that the initial reaction is monomolecular. We will refer later in this paper to the exact nature of this reaction; but it may be mentioned that many facts suggest that the initial process might be a change in the molecular architecture of glucose, brought about by adsorption to some specific enzymatic structure and prior to phosphorylation by adenosine-triphosphate. Such a reaction would indeed be of the monomolecular type; moreover, the process at this stage would still very probably be reversible, which would account for the phenomenon of "*stockage lacunaire*" observed by Soula and his co-workers. Nevertheless, too much significance should not be attributed to the linearity of the curve; as we have seen, it occurs only in a definite blood-sugar range and any limited portion of a parabolic curve is approximately a straight line. On the other hand, the fact that enzymatic processes are conditioned by adsorption to a surface renders kinetic deductions from the type of curve open to criticism.

The importance of this relationship in regard to insulin action is evident. It shows that the rate of blood-sugar level should be measured regularly, and the automatic influence of its variations should be taken into account (see later: The method of compensation (p. 43)).

Blood-sugar variations can affect the experimental results obtained with insulin in yet another way. In 1934 attention was drawn by Macleod (46) to the extraordinary differences existing between normal and diabetic animals in regard to their responses to insulin. While the metabolic changes brought about by insulin in diabetic animals are fairly regular and may be called "normal", i.e., opposite to the corresponding changes caused in normal animals by total extirpation of the pancreatic gland, the reactions of the normal animal to insulin are much more irregular. In a certain number of experiments, the response is also "normal," in others no changes occur, in many others, finally, "abnormal" diabetic reactions to insulin are observed, for instance a lowering of the R.Q., a decrease in hepatic glycogen, an increase in the production of ketone bodies, or a rise in nitrogen excretion, etc.

Macleod failed, however, to give a satisfactory explanation of these differences. There is no doubt that they are mainly due *to differences in blood-sugar level.* A systematic review of those earlier experiments, and of the later ones where similar differences occur, has shown us that most of the "abnormal" results were obtained on animals receiving large amounts of insulin and no glucose or insufficient quantities of glucose, while the "normal" results are observed on either diabetic animals, or on animals receiving only small doses of insulin, or adequately supplied with glucose whenever greater quantities of insulin were given. In other words, *"abnormal" results are generally associated with hypoglycaemia, "normal" ones with hyperglycaemic or normal blood-sugar levels.* In many cases, this relationship appears in the same set of experiments. From this fact may be concluded that the "abnormal" results are mostly due to the action, not of insulin itself, but of *blood-sugar raising or diabetogenic factors whose appearance in the blood-stream is induced by the hypoglycaemic state brought about by insulin.* For this reason

marked hypoglycaemia must be considered as the *chief cause of error* in experiments on insulin and should be avoided at all costs.

In short, the blood-sugar level appears to be of primary importance in all experiments concerned with insulin. We believe that too much stress cannot be laid on its importance, and that the confusion still existing amongst different schools, and the resulting failure of any theory of insulin action yet proposed to carry unanimous agreement, are to be attributed mainly to neglect of this factor, or insufficient knowledge of its influence. As has been shown, this influence is twofold:

1, the rate of glucose consumption is directly related to its concentration in the blood; 2, low blood-sugar levels may induce secondary reactions which antagonize the action of insulin.

b. *The so-called "glycogenolytic action" of insulin.* It has been frequently stated, especially by French and American authors, that insulin exerts a glycogenolytic influence on the liver. In fact, a decrease in liver-glycogen after injection of insulin has been frequently observed; but so, for that matter, has an increase. Of the two results, the former undoubtedly may be termed "abnormal" in view of the diminution in hepatic glycogen reserves resulting from pancreatectomy. In no other set of experiments does the "glycaemic criterion" prove to be such a valuable asset as in these: it may be shown that many of the experiments in which glycogenolysis occurs are associated with hypoglycaemia and are therefore without demonstrative power; it has even been shown in some experiments (see Reid (56), Crandall and Cherry (14)) that glycogenolysis *fails to occur* in adrenalectomized or adrenal-denervated animals treated with insulin, thereby clearly demonstrating that this effect is due, not to insulin action but to a secondary discharge of adrenalin, induced by insulin-hypoglycaemia.

In some cases however, we either have no positive proof that such a reaction has occurred, as no recording of the blood-sugar curve has been made (e.g., Cori and Cori (13)), or may even definitely eliminate the possibility of such a reaction, as, for instance, in the experiments by Bodo and Marks (2), Fiessinger *et al.* (29), Lundsgaard *et al.* (43), where a glycogenolytic action of insulin is demonstrated on *the perfused liver.* An interesting fact about the latter experiments is that this effect on the perfused liver can be produced *with certain samples of insulin only;* others, at least equally active therapeutically, exert no glycogenolytic action whatsoever on the perfused liver. Evidently, even in these cases, glycogenolysis cannot be due to insulin itself; it must be attributed to a blood-sugar raising factor, different from insulin, which is present as an impurity in certain samples of commercial and even crystalline (Fiessinger, Lundsgaard) insulin. It will be remembered that a similar explanation accounts satisfactorily for the initial hyperglycaemia which occurs immediately after injection of certain samples of insulin.

Our attention was particularly drawn to this problem by a paper by Bridge (8), where it is categorically stated that insulin inhibits glycogen formation in the liver. In Bridge's experiments, addition of varying amounts of insulin to the solution of glucose infused continuously to rabbits, during 6 hours and at such

a rate as to maintain the blood-sugar at a constant hyperglycaemic level, leads to a corresponding *decrease* in the amount of glucose deposited as liver-glycogen. The insulin used in this experiment was "*Lilly;*" since this brand has been widely used for research purposes, we thought it rather important to verify whether it also contains a blood-sugar raising factor, as Bridge's results strongly suggested. Our experiments fully confirmed this assumption (24) (26).

Injection of "*Lilly*" insulin produces a characteristic initial hyperglycaemia; less glucose is needed to compensate the hypoglycaemia induced by massive injection of 30 units of another brand of insulin ("*Novo*") to rabbits, when "*Lilly*" insulin is added to the glucose solution; on the contrary, addition of "*Novo*" insulin to the glucose infused leads to an *increase* in the glucose requirements of the animal. When sufficiently high doses of "*Lilly*" insulin are added (16 to 23 U/Kg./hr), *no glucose at all is needed;* the blood-sugar level never falls below 0.09 per cent. The glycogenolytic factor is then supplied at a rate sufficient to activate the hepatic output to such an extent as to replace almost completely a continuous glucose infusion of 1,40 g./Kg./hr. From the quantitative data obtained, it may be calculated that Bridge would have observed an *increase in hepatic glycogen, proportionate to the amount of insulin given*, if he had used another brand of insulin and if he had further maintained all his animals *at the same blood-sugar level*, by adequately regulating the glucose infusion.

We feel that we may safely conclude that insulin itself exerts no glycogenolytic effect whatsoever; on the contrary, provided sufficient glucose is supplied to maintain the blood-sugar at a normal level, hepatic glycogen will regularly be found to increase under the influence of insulin. Further objections which might be brought against this conclusion will be met later in this paper (see "The final fate of glucose disappearing under insulin action (p. 52)).

c. *The method of compensation.* Apart from the fact that the brand of insulin used must be tested as to the possible presence of a blood-sugar raising impurity, the most important factor which should be kept in mind in all experiments concerning insulin action is the level of the blood-sugar. Unless one wishes to study the reaction to hypoglycaemia, low blood-sugar levels must be prevented from occurring; moreover, the direct influence of blood-sugar concentration on blood-sugar disappearance must be taken into account. The most perfect method will therefore be the one in which the blood-sugar is continuously maintained at its normal level by an adequate infusion of glucose. Already in 1929 such a method was devised by Bouckaert, De Nayer and Krekels (4) (5), and it has been widely used subsequently by Bouckaert and his co-workers. Lately, de Duve and Bouckaert (19) have pointed out that the original method can be rendered more serviceable and more precise by taking advantage of the linear relationship existing between blood-sugar level and consumption. In both methods the insulin is given massively by intravenous injection at the beginning of the experiment and this injection is immediately followed by a continuous intravenous infusion of isotonic glucose; it can be shown that if the blood-sugar variations are not excessive, the quantity of glucose infused corresponds, with less than 5 per cent error, to the quantity which has effectively disappeared under insulin action and

provides a very accurate measure of this action. At least, this will be true if the independent and automatic influence of the blood-sugar level itself is taken into account. This was done in the original method by retaining only those experiments in which the blood-sugar curve constantly remained within "physiological limits"; a large number of experiments had thereby to be discarded and the method proved onerous and unpractical. In the new method, the mean blood-sugar level G is calculated from the blood-sugar curve, and D being the quantity of glucose given (i.e., consumed at a mean blood-sugar value G), the quantity D_0 *which should have been given* to maintain a normal blood-sugar concentration G_0 is calculated by the experimentally established formula:

$$D_0 = D \times \frac{G_0}{G}$$

Through this simple mathematical process, the range of blood-sugar curves which remains of practical use is far greater, and the measure of insulin action is far more accurate as all the results are reduced to an identical blood-sugar level G_0, the independent influence of the level of the blood-sugar being thereby eliminated. When assays of this type are performed on a set of similar animals, the values of D_0 usually differ from one animal to the other and can be represented by only a mean value. Finally, the mean value of D_0 thus obtained provides a quantitative test of the action of a given dose of insulin on a given type of animal.

Another modification of the original method has been devised by De Maeyer and De Graer (16). In this second method, the values of G are plotted against the values of D, and the most probable curve relating D and G is calculated. The intersection of this curve with a vertical line passing through the value of abscissa G_0 furnishes directly the most probable value of D_0. The two methods give almost identical results.

It is to be noted that both these methods imply the use of statistical formulae. It is very important, especially in biological research where the number of experiments is usually small, that statistical results should be presented properly, that is, accompanied by their dispersion. This can be done with the aid of the formulae given by Fisher (30) in his book, *Statistical methods for research workers* (see also (16) (19) (27)).

d. *The site of insulin action.* For more than 50 years, two theories have been opposed to explain the origin of diabetic hyperglycaemia and of insulin hypoglycaemia. According to one theory, the main site of insulin action is the liver and diabetic hyperglycaemia is mainly due to overproduction of glucose in this organ, while insulin hypoglycaemia is essentially a result of an inhibitory action on liver glycogenesis. According to the other, insulin acts mostly on the muscles, non-utilization of glucose being the cause of diabetic hyperglycaemia, excessive utilization that of insulin hypoglycaemia. In this second theory, it is assumed that the intense variations in hepatic metabolism which undoubtedly occur in diabetic animals are mainly due to secondary reactions provoked by the primary glucose need of the muscles.

Up to 1930, the muscular theory was advocated by a majority of authors and is still widely accepted today. The arguments for it are manifold: 1. While the fact that insulin acts directly on the glucose consumption of muscle tissue was firmly established by a great number of authors, none of the experiments carried out to reveal a similar action on the isolated liver had given conclusive results. 2. While, on the whole animal, changes in muscular metabolism were fairly regular after administration of insulin, hepatic responses to insulin proved excessively variable and were even often "abnormal," i.e. similar to those occuring after pancreatectomy. 3. Some well known experiments, namely those by Mann and Magath, Lesser and co-workers, Cori and Cori, Best, Dale, Hoet and Marks appeared particularly demonstrative and have been frequently quoted in support of the muscular theory.

Since then, the hepatic theory has gained ground thanks to the work of Soskin and his co-workers; nevertheless, none of Soskin's experiments conclusively disproves the hypothesis of an indirect action of insulin on the liver.

In order to measure quantitatively the amount of glucose which disappears respectively in the liver and in the peripheral tissues under the action of insulin, we have carried out comparative tests, using the method described above, on whole dogs, on hepatectomized dogs, eviscerated dogs and pancreatectomized-hepatectomized dogs, either receiving super-maximal doses of insulin or receiving no insulin at all (de Duve et al. (1945) (20)).

Our results show not only that insulin may inhibit completely hepatic glycogenesis in the animal with a normal blood-sugar but may even further lower the hepatic threshold, thereby considerably increasing glucose retention by the liver. They prove that insulin acts directly on the liver and that the hepatic action is more imporant than the peripheral action.

Here is a more detailed description of the procedure followed in these experiments:

Using the method described in the preceding chapter, we have ascertained that, to maintain a normal blood-sugar level in a dog constantly during 80 minutes after he had received a super-maximal dose of insulin given intravenously, it is necessary to provide, by continuous intravenous infusion, a total quantity of glucose equal to 2,10 gr./Kgr. From this result may be gathered that the various tissues of a dog saturated with insulin remove from the blood-stream large quantities of glucose corresponding to 2,10 gr./Kgr./80 min. provided the blood-sugar level itself does not vary. It will be remembered that if the rate of glucose infusion is such that variations of the blood sugar level occur, this would, under the same conditions of insulin excess, affect the quantity removed, this quantity being larger than 2,10 gr./Kgr. if the dog becomes hyperglycaemic, smaller if hypoglycaemia occurs. We may further assume that, in this experiment, the action of insulin is observed as much as possible uncomplicated by other factors, such as adrenalin and other blood-sugar raising hormones, whose secretion would be increased if the animal were allowed to become hypoglycaemic.

The point we wanted to investigate was the behaviour of the liver under those conditions. We know that the liver of the fasting dog secretes into the blood

stream a quantity of glucose covering the total consumption of the peripheral tissues. We know also of three factors which may alter this process: adrenalin will stimulate sugar formation in the liver by activating glycogenolysis; pituitary and cortical hormones will stimulate it also, mainly by promoting glyconeogenesis; finally, blood-sugar level variations will directly affect the hepatic glucose output, hypoglycaemia leading to its increase and hyperglycaemia bringing about an inhibition of this process, an inhibition which may be partial, or complete, or may even give place to the opposite process of retention, according to the intensity of the hyperglycaemia obtained. In our experiments, changes in blood-sugar level are avoided so far as this is possible; on the other hand, we have no reason to believe that important alterations take place in the secretion of the adrenal and pituitary glands, since their most powerful stimulant, which is the blood sugar level itself, suffers no great variations.

The most important factor obtaining under these experimental conditions will be therefore the considerable excess of insulin to which the animal is submitted, and these conditions appear particularly suitable for the study of the hepatic reaction to this hormone.

If we assume that insulin does not at all affect hepatic metabolism, we must expect the hepatic glucose output to continue undisturbed and the quantity of glucose thus poured into the bloodstream will add itself to that given by infusion to meet the requirements of the peripheral tissues; in that case the value of 2,10 gr./Kgr. should indicate the difference between the peripheral requirements of the animal saturated with insulin and those of the animal receiving physiological amounts of insulin from its own pancreas. If insulin does promote sugar formation in the liver, the excess of glucose thus provided would have to be added to the quantity infused, the sum of the three being equal to the peripheral requirements of the insulinized animal. On the other hand, if insulin partially inhibits sugar formation in the liver, the peripheral requirements of the insulinized animal will be less than the normal fasting glucose output plus 2,10 gr./Kgr., but they will still be larger than 2,10 gr./Kgr.; if insulin completely inhibits sugar formation in the liver, the peripheral requirements of the insulinized animal will be exactly 2,10 gr./Kgr.; and finally, if insulin not only inhibits sugar formation in the liver but further promotes hepatic retention of glucose, the peripheral requirements of the insulinized animal will be smaller than 2,10 gr./Kgr. by a quantity equal to that taken up by the liver under the influence of insulin.

It is clear than an adequate measurement of the *peripheral glucose requirements* of an animal saturated with insulin will allow us to distinguish between these various hypotheses. We have therefore determined the quantity of glucose which must be furnished by continuous infusion to maintain a constantly normal blood sugar level on a number of hepatectomized and totally eviscerated dogs receiving a super-maximal dose of insulin. We have also carried out similar determinations on hepatectomized, on totally eviscerated and on pancreatectomized-hepatectomized dogs, receiving no insulin.

The various results obtained are summarized in table 1.

The most significant results are those obtained on hepatectomized and evis-

cerated dogs receiving large amounts of insulin; they show that exclusion of the liver strongly diminishes the glucose requirements of the insulinized animal and enforce the conclusion that insulin given in large amounts to an animal whose blood-sugar level is kept normal by a continuous infusion of glucose, not only completely inhibits the hepatic glucose-output, but even considerably activates glucose retention by the liver, the amount of glucose taken up by the liver under those conditions (2,10 − 0,40 = 1,70 gr./Kgr./80 min.) being much larger than the total amount taken up by all the other tissues (0,40 gr./Kgr./80 min.).

A more detailed discussion of the results serves to emphasize this conclusion. The results obtained on the hepatectomized animals receiving no insulin give us an idea of the state existing in the normal fasting animal supplied with insulin from its own pancreas. Its peripheral glucose requirements are 0,32 gr./Kgr./80 min.; there is no indication that hepatectomy diminishes the glucose consumption

TABLE 1

NUMBER OF EXPERIMENTS	OPERATIVE DEPRIVATION	INSULIN GIVEN	INSULIN AVAILABLE	GLUCOSE REQUIREMENTS TO KEEP THE BSL CONSTANTLY AT ITS NORMAL VALUE OF 0,87 GR. P. L. GR./KGR./80 MIN. $\bar{x} \pm T\, s/\sqrt{n}$ *
7	None	+	Super-max.	2,10 ± 0,24
6	Hepatectomy	+	Super-max.	0,40 ± 0,12 } 0,40 ± 0,08
4	Evisceration	+	Super-max. and under-max.	0,40 ± 0,22
3	Hepatectomy	−	Norm. endogenous	0,32 ± 0,06
3	Evisceration	−	Remaining in tissues immediately after evisc.	0,18 ± 0,15
2	Previous pancreatectomy hepatectomy	−	None	0,09 and 0,14

* $\bar{x}$: mean. s: standard error of mean. n: number of experiments. T: "students value of t for P = 0.05 and n values. In less than 5 per cent of cases shall the true value of $\bar{x}$ fall outside the limits indicated by the value affected by the sign ± (Fisher p. 111).

of the peripheral tissues, since this figure is not smaller than those obtained by other authors, some of them on intact animals.[3]

We may take it, therefore, that the peripheral tissues of the normal fasting animal take up 0,32 gr./Kgr./80 min. of glucose from the blood-stream, the same quantity being poured into the blood-stream by the liver. In other words, the liver supplied with physiological amounts of insulin secretes 0,32 gr./Kgr./80 min. of glucose; under the influence of the excess of insulin injected, glucose retention by the liver passes from a negative figure of −0,32 to a positive one of +1,70/Kgr./80 min. the total change on balance produced being, therefore, 2,02 gr./Kgr./80 min. Under the influence of the same excess of insulin, peripheral consumption is raised from 0,32 to 0,40 gr./Kgr./80 min. the excess of

[3] On the contrary, there seems to be rather an increase in the peripheral consumption of glucose after hepatectomy. See p. 51–52.

glucose taken up being 0,08 gr./Kgr./80 min. It is clear that doses of insulin higher than the physiologically secreted amounts of this hormone act almost exclusively on the liver, where they promote glucose-retention in a remarkable way, the ratio-hepatic action/peripheral action being 25 (table 2).

On the other hand, the smaller physiological doses of insulin exert a strong action on the peripheral tissues. This is shown by the results obtained on eviscerated and on pancreatectomized-hepatectomized animals receiving no insulin. The smallest figure we have obtained for the peripheral consumption of an animal completely deprived of insulin is 0,09 gr./Kgr./80 min. It will be remembered that this figure has been reduced to that for a normal blood-sugar level; in reality, a consumption of 0,24 gr./Kgr./80 min. has been observed at a mean glycaemia of 0,239 per cent.

By comparing this figure with that observed on the hepatectomized animal supplied with insulin from its own pancreas (0,32 gr./Kgr./80 min.), we find that physiological amounts of insulin promote the peripheral disappearance of 0,23 gr./Kg./80 min. of glucose. The figure obtained on the eviscerated animal (acutely pancreatectomized-hepatectomized) is an intermediate one (0,18 gr./Kgr./80 min.), showing that the insulin remaining in the tissues immediately after pancreatectomy still favours glucose consumption to a certain extent.

TABLE 2

AMOUNTS OF INSULIN AVAILABLE	AMOUNTS OF GLUCOSE TAKEN UP AT A NORMAL BS LEVEL GR./KGR./80 MIN.		RATIO: HEPATIC ACTION/PERIPHERAL ACTION
	Liver	Other tissues	
Super-maximal.	+1,70	+0,40	
Physiological..	−0,32	+0,32	
Difference....	+2,02	+0,08	25

In order to determine the ratio: hepatic action/peripheral action of physiological amounts of insulin, we should know to what extent the insulin secreted by the pancreas inhibits sugar-formation in the liver of the fasting normal animal; in other words, we should know how much glucose would be poured into the bloodstream by the liver of an animal, completely deprived of insulin without the blood sugar level having changed. This figure cannot be obtained ex-experimentally; even without taking into consideration the change in blood sugar level, the hepatic glucose output of a diabetic animal would probably vary considerably according to its state of nutrition and to the degree of ketogenesis attained. In the most favourable conditions, such as in the protein-fed diabetic animal, the liver supplies large amounts of glucose to the blood stream, covering the peripheral needs (approximately normal) and the renal excretion. Taking into account the inhibition of the hepatic output produced by the high blood-sugar level, we have estimated that the liver of a dog, kept at a normal blood sugar level and completely deprived of insulin would secrete into the bloodstream

approximately 0,85 gr./Kgr./80 min. (de Duve (27)). Although this figure rests on only very indirect evidence, it may serve to emphasize an important fact, for there is little chance that the hepatic action of physiological amounts of insulin is more than that of reducing the hepatic output from 0,85 to 0,32 gr./Kgr./80 min., i.e., that of favouring the retention of more than 0,53 gr./Kgr./80 min. of glucose. And we may take it that the ratio: hepatic action/peripheral action of these small doses of insulin will hardly exceed the value of 2 (table 3).

The same reasoning applied to the total action of a super-maximal dose of insulin furnishes for the ratio: hepatic action/peripheral action, a value of approximately 8 (table 4).

TABLE 3

AMOUNTS OF INSULIN AVAILABLE	AMOUNTS OF GLUCOSE TAKEN UP AT A NORMAL BS LEVEL GR./KGR./80 MIN.		RATIO: HEPATIC ACTION/PERIPHERAL ACTION
	Liver	Other tissues	
Physiological	−0,32	+0,32	
None	−0,85	+0,09	
Difference	+0,53	+0,23	2

TABLE 4

AMOUNTS OF INSULIN AVAILABLE	AMOUNTS OF GLUCOSE TAKEN UP AT A NORMAL BS LEVEL GR./KGR./80 MIN.		RATIO: HEPATIC ACTION/PERIPHERAL ACTION
	Liver	Other tissues	
Super-maximal	+1,70	+0,40	
None	−0,85	+0,09	
Difference	+2,55	+0,31	8

Finally, we may conclude:

1. That the higher the dose of insulin, the more it tends to act on the liver; the *maximal active dose* must therefore be much higher for the liver than for the other tissues.

2. That, whatever the dose of insulin present, and especially with the larger ones, much more glucose is taken up by the liver than by the other tissues under the influence of insulin; consequently the *maximal amount of glucose* which can be taken up under the influence of insulin at a normal blood-sugar level is also much larger for the liver than for all the other tissues.

Taken together, these facts prove that insulin exerts a direct and powerful action on the liver and that this action is far more important, quantitatively, than its peripheral action.

These results are in full agreement with those of Soskin's and, added to them, should afford convincing proof of the essential and central rôle played by the liver

in the processes of blood-sugar regulation. They raise many questions and lead to many conclusions, of which we can only, and very briefly, mention the most important ones.

One may wonder, first of all, how this important action has escaped the attention of so many authors and has even been emphatically denied by some. Two facts are ma nly responsible for this: 1. As we have already mentioned above, neglect of the importance of the blood-sugar level has led to many erroneous results, particularly in those experiments where hypoglycaemia has not been avoided. It is clear that the antagonistic reactions occurring in hypoglycaemic animals consist essentially in a discharge of hormones which stimulate the glucose production of the liver and thereby electively oppose the hepatic action of insulin without interfering with its peripheral action. The use of impure insulin, contaminated by a blood-sugar raising factor, has also furnished faulty results. 2. The hepatic systems on which insulin acts are extremely fragile. We were able to note this fragility in a number of experiments (21) (23), in which we observed a deleterious effect on insulin action of various factors (adrenalin, anesthetics, hyperventilation, hypoventilation, spinalisation). Of all these factors, reduction of the oxygen-supply and spinalisation were the most noxious. The fact that insulin is inactive on the perfused liver is therefore not astonishing, especially if it be remembered that the liver is particularly well supplied in enzymes of all sorts and for that reason, very sensitive to autolytic influences. It may be further added that our experiments do not allow us to exclude the possibility that insulin, as has been suggested by Nielsen (52), must be activated by a substance of muscular origin, to be active on the liver.

If a complete survey of the earlier literature were made with these fact ins mind, it would be clear that a feeling of distrust toward the existence of a hepatic action of insulin was created not so much by lack of evidence as by the confusing amount of contradictory results obtained in this field. Once these contradictions are satisfactorily accounted for, the evidence in favour of the hepatic theory becomes overwhelming.

It may be objected to our experiments that the figures for peripheral glucose consumption obtained on anesthetized and artificially ventilated animals, submitted to severe operative interference, cannot be applied as such to the normal animal. This objection is easily met: 1. Our results are in close agreement with those obtained by various other authors who have carried out similar experiments, showing that the technique used in our experiments is not inferior to that used by other authors. 2. As far as the validity of the technique itself is concerned, an interesting fact emerges from a comparison of our results with those of Cherry and Crandall (10) and of Soskin, Essex, Herrick, and Mann (59). Cherry and Crandall (10) have measured the hepatic glucose output of normal unanesthetized dogs, by means of the angiostomy method; they find that the liver secretes 9,1 $\pm$ 0,68 mgr. of glucose p. 100 ml. of hepatic blood flow. According to the results of Blalock and Mason (1), the mean hepatic blood flow of the normal dog is 28,6 ml./Kgr./min. From these two figures it may be calculated that the mean hepatic glucose output of a normal unanesthetized dog, and consequently also

his mean peripheral glucose consumption, amounts to 0,21 gr. ± 0,02/Kgr./80 min. Using a similar method Soskin, Essex, Herrick and Mann (59) have measured simultaneously the hepatic glucose output and the hepatic bloodflow of anesthetized dogs, before and after administration of varying amounts of glucose. If we single out in their experiments the 6 figures obtained on dogs with a normal blood-sugar level, we find a mean value of 0,21 ± 0,12 gr./Kgr./80 min. There is no indication, therefore, that the anesthesia has greatly modified either the hepatic output or the peripheral consumption in these animals. Finally, we ourselves have found for the peripheral consumption of anesthetized, hepatectomized dogs a mean value of 0,32 ± 0,06 gr./Kgr./80 min. Neither the anesthesia nor the exclusion of the liver seem to have brought about any diminution of the capacity of the peripheral tissues to take up blood glucose. On the contrary, the figure found on the hepatectomized dog appears to be significantly higher than that observed on the intact animal. The obvious explanation of this difference seems to be that, in the hepatectomized animal, all the insulin secreted by the pancreas is carried out directly to the peripheral tissues, whereas in the intact animal, part of it is retained by the liver. This view is supported by the fact that, in the totally eviscerated dog, the peripheral consumption is much smaller, approaching the value observed on the intact animal (0,18 gr./Kgr./80 min.).

In the latter case no insulin at all is conveyed to the tissues during the experiment, but, considering the experiment is carried out immediately after the operation is performed, the effect of the insulin present in the tissues must be expected to wear off only gradually, which accounts for the fact that the glucose uptake of the eviscerated dog appears to be only slightly lower than the peripheral consumption of the intact animal.

3. Finally, it may be noted that both the hepatectomized and the eviscerated dog are perfectly sensitive to their own insulin as well as to externally provided insulin.

All these facts seem to deal satisfactorily with the objection that the important reduction of the capacity to utilize glucose under the influence of insulin, brought about by extirpation of the liver, could be due to a depressing influence of the operative interferences on the peripheral capacity of utilization. On the other hand, the various quantitative data by which we have tried to systematize our results may have to be revised in view of the fact that hepatectomy apparently induces a state of relative hyperinsulinemia in the peripheral tissues. If we take as a more probable value for the peripheral consumption of a normal dog, the figure of 0,21 gr./Kgr./80 min., observed on intact animals both by Cherry and Crandall and by Soskin *et al.*, the figures given above should be changed to the following values:

Hepatic retention with physiological amounts of insulin: −0,21
Peripheral consumption with physiological amounts of insulin: +0,21
Hepatic retention with complete lack of insulin: approximately −0,69.

In that case, the final figures given in tables 2, 3 and 4 should be replaced as follows (table 5).

It is clear that these corrections do not affect at all our main conclusions. In-

deed, they appear to provide corroborative evidence in favour of a prevalent action of insulin on the liver. From the comparison which has been drawn between our results and those of Cherry and Crandall and of Soskin et al., it appears that, through extirpation of the liver, the amount of glucose disappearing in the peripheral tissues under the action of physiologically secreted insulin is almost doubled. As the relation between the amount of glucose consumed under the influence of insulin and the amount of insulin present in the animal is represented by a typical saturation curve (Bouckaert, De Nayer and Krekels (4) (5)), it follows that the amount of insulin made available to the peripheral tissues by hepatectomy is more than doubled; in other words, that the liver normally retains the largest part of the insulin secreted by the pancreas. As has been shown, this fraction is not merely retained by the liver to be destroyed, but exerts a powerful action on this organ. The possibility remains, all the same, that the liver may also intervene in regulating the blood-sugar level, by allowing more or less insulin to reach the peripheral tissues.

e. *The final fate of glucose disappearing under insulin action.* A general answer to this question can easily be drawn from the results obtained in our laboratory

TABLE 5

TABLE	AMOUNTS OF INSULIN ACTIVE	GLUCOSE DISAPPEARING UNDER THE INFLUENCE OF INSULIN IN		RATIO: HEPATIC ACTION/PERIPHERAL ACTION
		Liver	Peripheral tissues	
II	Excess (maximal − physiological)	+1,91	+0,19	10
III	Physiological	+0,48	+0,12	4
IV	Total maximal dose	+2,39	+0,31	8

by Bouckaert and Stricker (3) in 1924 and by Lamers (38) in 1926. Bouckaert and Stricker have measured by direct calorimetry the heat production of rabbits receiving glucose and insulin and observe a maximum rise of 10 per cent. Lamers, using indirect calorimetry, has measured the total heat-production and the respiratory quotient of rabbits in the same conditions. According to his results, the basal metabolism suffers only slight variations but there occurs a definite rise in R.Q., approaching unity. From these data, the total amount of carbohydrate oxidized can be derived.

The experiments were performed before the method of compensation was standardized; in some the glucose was administered subcutaneously, in others intravenously but a precise measure of the quantity which had disappeared was not available. Luckily, in a number of those experiments, super-maximal doses of insulin were used and the glycaemic curve is contained within physiological limits; from the data obtained subsequently by Bouckaert, De Nayer and Krekels (4) (5), we know the exact amount of glucose which has disappeared under those conditions and can now compare this amount with the amount oxidized.

This comparison has been done recently in a paper by Bouckaert and De Nayer (1945) (7) and leads to the conclusion that only 50 per cent of the total

amount of glucose taken up in all the tissues has been oxidized. We must assume therefore that the other half has been stored under form of reserve substances, probably by oxidative synthesis. But, as the rise in glucose oxidation is not associated with a corresponding rise in total metabolism, it follows that an equivalent amount of protein and fat must have been spared as a result of the oxidation of the first half.

If we assume that a quantity 2 A has disappeared under insulin action, the final fate of this quantity may be gathered from the following data: A will be oxidized but the energy yielded by this reaction will be stored as chemical energy by the sparing of an equivalent amount of reserve substances; a further quantity α A will also be oxidized and the energy yielded will serve for the conversion of the remaining quantity $(1 - \alpha)$ A into reserve substances. Therefore, for every loss of glucose there is a corresponding gain of reserve substances, either spared or newly formed, and we may conclude that *when glucose disappears under insulin action the main result is the conservation and possibly the synthesis of reserve substances.*

A more detailed description of the processes involved raises four questions:

1. What are the reserve substances, which are spared or synthesized as a result of glucose disappearance under insulin action? The most important one is certainly glycogen. From the value of the R.Q. of the fasting animal, we deduce that part of its caloric requirements is furnished by direct oxidation of carbohydrate; this carbohydrate is almost exclusively present under the form of glycogen. When glucose is given along with insulin, this part of the caloric requirements is now met by the oxidation of glucose and an equivalent part of glycogen is spared. Moreover, many experiments show that glycogen is also newly formed when sufficient quantities of glucose and insulin are given. But the sparing action does not stop at glycogen and affects also fat and protein oxidation; otherwise a rise in R.Q. would not occur. And there is good evidence that, in certain conditions, fat and protein will not only be spared, but also be newly formed, when great amounts of glucose disappear under insulin action.

2. Do these results fit in with an essentially hepatic theory of insulin action? They definitely do so, for it will be remembered that the normal peripheral requirements are met mainly by glucose present in the blood. This glucose is provided, in the fasting animal, by formation from glycogen, protein and possibly fat, which takes place in the liver. Therefore the sparing action affects chiefly hepatic metabolism, where glycogenolysis and glyconeogenesis will be strongly inhibited, and it follows naturally that the main site of synthetic processes will also be the liver. Of course, peripheral metabolism will be influenced as well, especially by the smaller doses of insulin: this action will affect muscular glycogen reserves and peripheral stores of fat and protein.

3. Does the great amount of experimental work devoted to this subject support this theory? As we have pointed out before, many of the results obtained by different authors are contradictory, so that, taken together, they furnish insufficient evidence in favour of any theory of insulin action. It can be shown, however, that these contradictions are only apparent and are mainly due to differ-

ences of technique leading to erroneous interpretations. We can only deal very briefly with this important aspect of the problem, referring for further details to the work of de Duve (27), already mentioned, in which the former literature has been critically examined and reinterpreted.

It will be clear by now that the blood-sugar level plays the most important part in this reinterpretation. First of all, consideration of its importance allows the discarding of all experiments where marked hypoglycaemia has been provoked by insulin, insufficiently compensated with glucose, for the results obtained in those experiments are partly due to hormones antagonistic to insulin, such as adrenalin, diabetogenic pituitary factors, cortin, etc. This group will be found to comprise most of the experiments in which "diabetic" phenomena have been noted to occur after injection of insulin or in which the normal effects of insulin have failed to occur, for instance: those in which either no rise or a lowering of the R.Q. has been noted, those in which either no decrease or an increase in nitrogen excretion, or in ketogenesis and fat catabolism, have been observed, etc. Decreases in liver glycogen also chiefly occur in hypoglycaemic animals; this question has already been dealt with (see "The so-called glycogenolytic action of insulin", p. 42).

The question of blood-sugar level also intervenes in the interpretation of a second group of experiments, namely, those in which equal quantities of glucose are given to animals, one half of which are subjected additionally to an injection of insulin. Comparison of the results obtained on the insulinized animals with those observed on the non-insulinized animals frequently enforces the conclusion that *addition of insulin does not lead* to a greater formation of this or that substance than is produced by the glucose itself. Curiously enough, most authors have failed to note the important significance of this conclusion. If both the insulinized and the non-insulinized animal receive the same quantity of glucose, it is clear that in both animals the same quantity of glucose will have finally disappeared from the blood-stream into the tissues. If therefore a greater quantity of it is not found under one specific form, this only shows that *insulin does not favour one path of disappearance against the other, but treats all paths of disappearance alike.* The one great difference between the insulinized animal and the non-insulinized animal lies in the *difference in blood-sugar level* at which the disappearance of the same quantity of glucose takes place along the same paths. If the same experiment were performed in similar conditions of glycaemia, a greater quantity of glucose should have to be injected to the insulinized animal to maintain a blood-sugar level identical to that of the non-insulinized animal; the automatic influence of the glycaemic difference would then be eliminated and an increase in the formation of the substance studied with respect to the non-insulinized animal would indeed be noted.

Such results are extremely interesting, because they tend to show that *the fate of glucose disappearing under insulin action is not different from the fate of glucose disappearing under the automatic influence of an increase of its concentration in the blood, whatever the quantity of insulin available.* This fact is confirmed by the many experiments, in which the effects of glucose itself have been studied; it is

well known that provoked hyperglycaemia leads to an increase of the R.Q. and of hepatic and muscular glycogen reserves, exerts a sparing action on fat and protein catabolism and even furthers the synthesis of those substances. All these actions can be also attributed to insulin. Further confirmation is furnished by the experiments carried out on diabetic animals and showing that all the metabolic processes known to be furthered by insulin can be promoted without insulin being present, provided the blood-sugar level is sufficiently raised.

It becomes easily understandable hereafter why experiments performed on hypoglycaemic animals have furnished such misleading results. In a first stage of these experiments, insulin will promote the disappearance of the small quantities of glucose present in the blood-stream, creating hypoglycaemia; in a second stage, a stimulation of hepatic glycogenesis will take place under the influence of the blood-sugar raising factors secreted through reflex reactions against hypoglycaemia; glucose will then be secreted into the blood-stream and will be further metabolized under the influence of insulin in the tissues not involved in the antagonistic reactions, that is: mainly in the extra-hepatic tissues; the final result will be an elective furthering of peripheral consumption, frequently combined with activation of hepatic glycogenesis. The results might have been different if insulin had proved to be an activator of one specific reaction of glucose metabolism, such as glycogen synthesis, for instance, but we have seen that this is not the case.

Once these important facts are made clear, the interpretation of apparently conflicting results presents little difficulty and all experiments which are not open to similar criticism are found to agree with the theory expounded above. In a few cases, the problem is less simple and some results remain difficult to explain. Some of these are due, as has been shown above, to the presence of a glycogenolytic substance in certain samples of insulin. Another instance of them is furnished by the recent results of Britton and Corey (9) (12), in which it is stated that cortin, but not insulin, promotes the synthesis of hepatic glycogen. We have expounded elsewhere why we do not think their results with insulin to be conclusive (24) (27) and only wish here to emphasize the fact that cortin, as well as anterior pituitary extracts, further glycogen synthesis by furthering glyconeogenesis, and perhaps also by activating phosphorylase (Verzar), while insulin promotes glycogen formation in the liver by increasing the total amount of glucose taken up by the liver from the blood-stream. As a result of this, an increase in hepatic glycogen will be best obtained by the combined action of insulin and of cortin or anterior pituitary extract, for glycogen will be the only form under which the glucose taken up by the liver from the blood-stream as well as that produced in the liver from protein and fat can be stored. This has been conclusively proved by Marks and Young (47) and by Loubatières (42). But there is no reason why excessive stimulation of glyconeogenesis (by cortin) should not further glycogen formation either in the liver of pancreatectomized animals or in the perfused liver; on the other hand, inhibition of glyconeogenesis (by corticectomy) may alter the fate of glucose taken up by the liver under insulin action by favouring the paths leading to protein and fat against that leading

to glycogen; this will particularly be true if Verzar's theory be confirmed, as glycogen synthesis would then itself be impaired.

4. Finally one may wonder how insulin promotes glucose disappearance and creates such complex metabolic changes involving fat and protein as well as carbohydrate metabolism. To answer this question, it is necessary to outline briefly the theory of insulin action which will be further expounded at the end of this paper.

This theory is based essentially on the *analogy existing between the effects produced either by an increase in the blood-sugar level without addition of insulin or by a discharge of insulin without variation of the blood-sugar level*

In both cases, as is shown by a comparison of the results of Bouckaert, De Nayer and Krekels (4) (5) with those of de Duve and Bouckaert (19), a surplus of glucose disappears from the blood-stream. In both cases, as may be inferred from our own results (20) compared with those obtained by Soskin, Allweiss and Cohn (57), Soskin, Essex, Herrick and Mann (59), the main site of this disappearance is the liver, only a relatively small fraction of the total amount disappearing in the extra-hepatic tissues. In both cases finally, as it has just been shown, the final fate of the glucose disappearing is the same.

The question stated above can therefore best be answered by identifying first the mechanism by which the concentration of glucose itself promotes its own disappearance. It will be remembered that this relationship is independent of all neuro-hormonal control and must be due to the biochemical properties of the tissues involved. The law of mass action, applied to the initial reaction by which glucose enters the metabolic processes, provides an adequate explanation of this phenomenon. Indeed, this initial reaction may be represented by the equation:

$$\text{Glucose } (+ \text{ A } \cdots) \rightarrow \text{X}$$

and its velocity will be increased by an increase in glucose concentration. But this will lead to an increase in the concentration of X, which in its turn will increase the velocity of its disappearance, and so on. By virtue of the law of mass-action, the increase in glucose concentration will finally accelerate the whole chain of metabolic reactions conditioning its disappearance. Moreover its final fate will be different in different tissues, according to the biochemical potentialities of each tissue. Muscle tissue, for instance, will mainly oxidise glucose and convert it into glycogen, fatty tissue will synthesize fat out of it, etc.

In addition to this, the sparing action of glucose on protein and fat must also be explained. A possible explanation may be furnished by the fact that glucose and its immediate derivates have greater reducing powers than most metabolic substances and are therefore more easily oxidizable.

In the presence of a limited supply of oxygen, the oxidation of glucose would, for that reason, be privileged against that of other substances, and the sparing action of glucose would be nothing but a consequence of the fact that, in most cases, fat and protein are oxidized in the cells only when oxygen is rendered available by relative scarcity of glucose. Such an interpretation, coupled with the notion of "*biochemical inertia*" emphasized by Soskin and Mirsky (58) will be found to afford a satisfactory explanation of the gradual shifting of metabolism

to protein and fat, brought about by inanition or low carbohydrate diets, and of the enhancing effect of rich carbohydrate diets on glucose tolerance. Variations of the hormonal balance are of course also involved in these processes; what we wish to point out is that they are not the only factor responsible for them.

If this interpretation of the automatic effect of glucose concentration on glucose consumption be accepted, it follows necessarily from the similarity between this effect and the action of insulin, that insulin must produce the same primary results as an increase of the blood sugar level. In other words, insulin must also accelerate the initial reaction by which glucose enters the metabolic processes and must therefore be an activator of the enzymes responsible for that reaction. Such an explanation would account for all the phenomena[4] which are to be attributed to insulin action. In particular, it would explain how *insulin allows glucose to exert its multiple effects on hepatic and peripheral metabolism at lower concentrations than would otherwise be necessary.*

This theory can be tested by examining whether the substance X of the equation is effectively formed in greater quantity under the influence of insulin; in other words, by determining the immediate fate of glucose disappearing under insulin action.

f. *The immediate fate of glucose disappearing under insulin action.*[5] We have intentionally represented by the symbol X the first substance into which glucose is converted in the tissues, because the nature of this substance is not known with certainty. According to one school, led by Parnas, it is glycogen; other authors, with Meyerhof, identify it with hexose-monophosphate. Both alternatives have been considered in our laboratory in connection with insulin action. In 1932, De Nayer (17) determined the amount of glycogen present in two homologous muscles of rabbits, whose n. ischiatici had been previously cut to prevent muscular contractions, one muscle being taken before, the other after an injection of insulin adequately compensated by a continuous infusion of glucose. He has found that, although super-maximal doses of insulin were given, there was no increase but also no decrease (sparing action) in muscle glycogen, as long as the

[4] But one. See later: i) The action of insulin on protein metabolism, p. 66.

[5] The authors wish to make it clear that when this paper was sent in for publication, the later American literature was not yet available in Belgium. Even though some of the facts and theories discussed in these later chapters may now appear obsolete in the light of the recent findings on glucose phosphorylation by the Coris and their co-workers, they have preferred not to alter the original text in view of the material difficulties it would involve, and to go on stating the case as they were able to make it out under the conditions of enforced isolation imposed by the war.

Mention must at least be made of the recent announcement made by Price, Colowick and Cori (J. Biol. Chem. **160**: 633, 1945) that insulin is capable of reactivating "in vivo" and "in vitro" muscle hexokinase previously inhibited by anterior pituitary extract. Although this important observation affords only a partial explanation of the action of insulin, it constitutes the first definite step towards the elucidation of its biochemical mechanism. It also confirms the conclusion drawn from our own results that insulin must act as an activator of hexokinase. As to the nature of this activation and of the part played by hexokinase, the new evidence brings nothing either to prove or disprove the tentative explanation put forward in this paper.

blood-sugar level remained normal. Only by increasing the amount of glucose given and creating hyperglycaemia, did he succeed in promoting further glycogen synthesis in the muscles. Here again, the importance of the blood sugar level and the superiority of the method of compensation show quite clearly, when the results obtained by De Nayer are contrasted with those achieved by other authors in similar experiments.

Lately, de Duve and Hers (22) (25) have performed almost identical experiments, this time determining the amount of hexose-monophosphate. In the latter experiments, further precautions were taken and a control set of determinations was made on animals suffering the same treatment, but receiving an infusion of isotonic NaCl instead of glucose and insulin. A definite rise in the content of muscle hexose-monophosphate was observed in the insulinized animals, corresponding to a major part of the glucose taken up by the periphery.

This experiment, compared with that by De Nayer, may be termed crucial, both in regard to glucose metabolism and to the problem of insulin action. It shows that the substance X is hexose-monosphosphate, at least for the muscles, thereby confirming Meyerhof's theory, and it proves that insulin promotes the reaction leading from glucose to hexose-monphosphate, thereby confirming our theory of insulin action. The identity between the action of insulin and the automatic effect of hyperglycaemia is further demonstrated by comparing these results with those recently obtained by Leipert and Kellersmann (40), who have shown that infusion of excess glucose leads to an increase of muscle hexose-monophosphate, without affecting its content in glycogen.

We have not effected a similar test on the liver, the reason being that comparative determinations cannot be made on the liver of the same animal, without disturbing the glycaemic balance. But there are good reasons to believe that the uptake of glucose by the liver also occurs through phosphorylation. We know, for instance, that adrenalin strongly antagonizes the action of insulin and it has been shown that adrenalin acts on phosphorylative processes (see p. 68). According to Leipert and Kellersmann (40), absorption of glucose increases the content in esterified phosphorus of the liver; there is a good chance, consequently, that insulin does the same. Colowick, Kalckar and Cori (1941) (11) have shown glucose phosphorylation to take place in hepatic extract and have made a detailed study of this reaction (see p. 60). Finally, recent experiments by Kaplan and Greenberg (1944) (33) (34) (35) show that glucose tolerance is directly related to the amount of adenosine-triphosphate present in the liver and that both glucose and insulin increase this amount. As will be seen, adenosine-triphosphate plays the leading part in glucose phosphorylation.

It can reasonably be assumed, therefore, that the immediate fate of glucose disappearing under insulin action is to be converted into hexose-monophosphate, in the liver as well as in the muscles and other tissues.

g. *General theory of insulin action.* It may be useful, at this stage, to regroup the various facts discussed until now and to show briefly how they serve to support the theory of insulin action put forward in this paper. This section may be considered as a summary of the previous ones.

I. The most important fact which must be kept in mind is the important part played in blood-sugar regulatory processes by the blood-sugar level itself. The blood-sugar level regulates glucose consumption and production in the tissues both directly and indirectly.

1. By virtue of the law of mass action, the velocity of the reactions which condition glucose consumption is increased by a rise and decreased by a fall in blood-sugar level. Consequently, blood-sugar consumption is directly related to the blood-sugar level, independently of all neuro-hormonal processes, and this relationship has been shown to be linear within a range of glycaemia not too far removed from its normal value.

2. Hyperglycaemia provokes blood-sugar lowering neuro-hormonal reactions; hypoglycaemia stimulates blood-sugar raising neuro-hormonal factors.

II. Application of these principles to experiments concerned with the action of insulin leads to the important conclusion that variations in the blood-sugar level may considerably affect experimental results:

1. Hyperglycaemia may simulate the action of insulin, by automatically promoting glucose disappearance.

2. Hypoglycaemia will strongly antagonize the action of insulin, not only by the resulting automatic decrease in glucose consumption, but also by stimulation of antagonistic factors such as the orthosympathic system, adrenalin, cortin, anterior pituitary hormones, etc.

Consequently, all experiments where marked hypoglycaemia occurs should be discarded, as the results furnished by them are compound actions of insulin and of antagonistic factors. Moreover, in all experiments which are not liable to such criticism, the aùtomatic influence of blood-sugar variations must be taken into account, either by using a method in which all the animals are maintained at the same, preferably normal, blood-sugar level by a continuous infusion of glucose, or by reducing mathematically the results obtained to the same glycaemic level. Both these methods have been described and extensively used by us.

III. As a further precaution, the brand of insulin used in the experiment must be tested for the presence of a blood-sugar raising impurity, endowed with strong glycogenolytic properties. As we have been able to show, the use of such like insulin may lead to gross experimental errors.

IV. By using a method in which all these conditions are fulfilled, we have been able to demonstrate conclusively the existence and the primary importance of the hepatic action of insulin. Whatever the quantity of insulin administered, considerably more glucose disappears in the liver than in the other tissues under its influence. An excess of insulin which does not further promote peripheral glucose consumption, is still capable of furthering glucose retention by the liver.

V. The great number of conflicting results obtained with insulin is mainly due to neglect of the blood-sugar level. Proper reinterpretation of those experiments eliminates most discrepancies and allows the fate of glucose disappearing under insulin action to be accounted for satisfactorily. It becomes quite clear that, of the actions investigated, insulin produces no metabolic changes other than those

produced by the glucose itself: a rise in R.Q. occurs, associated with a negligible rise in total metabolism, thereby showing that if more glucose is oxidized, it is only utilized instead of other substances, glycogen already stored, fat and protein; when great amounts of glucose disappear, this sparing action is prolonged by new formation of glycogen, fat and protein. On the whole, when an excess of glucose disappears under insulin action, this glucose is stored as equivalent glycogen, fat and protein, either spared or newly formed. The real action of insulin is to allow glucose to exert similar actions at lower blood-sugar levels than would otherwise be necessary.

VI. The proximate fate of glucose disappearing in the tissues, whether under the action of insulin or not, is conversion into hexose-monophosphate.

VII. There is a strong analogy between the effects of an injection of insulin associated with a normal blood-sugar level and those of an increase in the blood-sugar level without change in the insulin supply. In both cases, an excess of glucose disappears, mainly in the liver, secondarily in the muscles and other tissues, is converted into hexose-monophosphate, which, in its turn, is partly oxidized, thereby sparing an equivalent quantity of glycogen, fat and protein, partly converted by oxidative synthesis, into glycogen, fat and, even to a certain extent, protein.

VIII. The effects produced by an increase in the concentration of glucose can only be explained through the fact that, by virtue of the law of mass-action, more glucose is converted into hexose-monophosphate. All the metabolic changes observed may be considered as necessary consequences of this primary change, the two main factors involved being a successive furthering of the whole chain of metabolic reactions concerned in glucose metabolism (mass effect), and the privileged oxidation of hexose-monophosphate and its derivatives, as compared with less oxidizable substances, such as fat and protein.

IX. We are finally led to conclude that insulin must act on the same primary reaction and in the same way as an increase in the concentration of glucose, and must therefore increase the velocity of this reaction. Insulin thereby appears as an activator of the enzymatic system concerned with glucose phosphorylation in the tissues, and this action of insulin is in itself sufficient to account for all the metabolic changes which appear to be a consequence of its influence (except one: see later i, p. 66).

h. *The biochemical mechanism of insulin action.* So far, we have been able to trace down the biochemical point of attack of insulin to the enzymatic system responsible for glucose phosphorylation in the tissues. If we want to probe further into the intimate mechanism of insulin action, our first concern must be to gather more particulars about the mechanism of glucose phosphorylation itself on one side, and about the biochemical properties of insulin on the other side, and to see whether and how both can be reconciled satisfactorily.

Luckily, the phosphorylation of glucose, together with other important phosphocatalytic processes, has been extensively studied, especially in later years. It was first seen to take place in yeast cells, but its intervention in the metabolism of mammalian tissues was long in doubt; more recently, however, the occurrence

of direct esterification of glucose into hexose-phosphate has been observed by a number of authors on a great variety of animal tissues and even in cell-free extracts of liver, brain, and heart-tissue; it has been subjected to a close study particularly by the Coris and their co-workers (Colowick, Kalckar and Cori (1941) (11)).

The findings made on yeast-cells and on animal tissues furnish closely similar results and enable a provisional picture of this reaction to be drawn.

There is no doubt that the main reaction consists in an exchange of phosphate between adenosine-triphosphate (ATP) and glucose, through which glucose is converted into hexose-monophosphate, ATP into adenosine-diphosphate or adenylic acid. But the complete process is more complex.

Already the early experiments of Meyerhof on yeast-metabolism have shown that yeast-extract loses very quickly the capacity of utilizing glucose, although it shows no loss in its power to metabolize starch or glycogen. Similar findings have been made on mammalian tissues. Isolated muscle or liver tissue becomes rapidly inactive towards glucose, while retaining its capacity to break down glycogen; muscle-extract, though endowed with unimpaired glycogenolytic properties, does not in the least attack glucose; and it is only lately that it has become possible to prepare cell-free extracts of certain organs, capable of phosphorylating glucose "in vitro."

All these findings tend to show that glucose cannot be phosphorylated as such, but must first be brought into an active state, before becoming vulnerable to the phosphorylating system; they also seem to indicate that the systems responsible for this activation are of a very fragile nature. This assumption has been confirmed by Meyerhof, who has isolated from yeast-cells a system capable of re-activating inactive yeast-extract and to which he has given the name "*hexokinase;*" a similar system has been shown to exist in fresh muscle-extract by Ahlgren, who has christened it "*glycomutin;*" an analogue of hexokinase has been isolated lately from brain-tissue by Ochoa (53). Nevertheless, no explanation has been yet proffered as to the nature of this "activation", which the molecule of glucose is supposed to undergo.

The following hypothesis, based on the reactions conditioning glycogen break-down in the muscles, has been put forward by one of us (27) to account for it: It will be remembered that the first reaction occurring in glycogen breakdown is a fixation of inorganic phosphate on glycogen (phosphorolysis), giving birth to glucopyranose—I. phosphate (Cori-ester). Hexose-monophosphate is further transformed into hexose-diphosphate by receiving one phosphate group from ATP (phosphorylation). This second reaction is very similar to the one through which glucose is phosphorylated, and it is to be noted that Cori-ester, which has the same structure as blood-glucopyranose, is no more able than glucose to react as such with ATP. It must previously be converted into fructofuranose-6-phosphate (Neuberg-ester). We may therefore, with Parnas (55), define the conditions of vulnerability of hexose-monophosphate to ATP as the following: the negative group must be attached to carbon 6 and carbon 1 must be set free under the form of a primary alcohol function by a changing of the pyranic into a

furanic cycle. It may be reasonably assumed that similar changes will have to take place inside the molecule of glucopyranose, prior to its phosphorylation by ATP; in that case, active glucose would prove to be fructofuranose, the action of hexokinase or glycomutin would be a temporary fixation of glucose by its carbon 6 associated to a regrouping of its pyranic cycle into a furanic one, and the first product of glucose phosphorylation would be fructo-furanose-1-phosphate (Robison-Tanko ester). It is to be noted that this ester has recently been isolated from the liver of rabbits by Pany (54). This hypothesis would explain also why fructose is more easily phosphorylated than glucose, and would account for the fact that the presence of an active form of glucose in the blood has been inferred mainly from experiments showing that the optical activity of blood glucose was inferior to the value calculated from its reducing power.

Activation of glucose is not the only condition necessary to render glucose phosphorylation possible; all the authors who have studied this reaction agree on the necessity for simultaneous oxidative processes, acting on a variety of substrates, amongst which pyruvic acid seems to be the most prominent one. Cozymase appears to play an important part in this oxidation. One may wonder with what particular stage of the reaction the oxidation of pyruvic acid is coupled. It is very improbable that it would be phosphorylation itself, as splitting of ATP is highly exothermic in itself. Two alternatives remain, which have been advocated:

According to Kaplan and Greenberg (1944) (33) (34) (35), the energy supplied by the oxidative process is used for the building up of ATP. The rate of phosphorylation of glucose being directly related to the amount of adenosine-phosphates available, it would be indirectly dependent on the intensity of the oxidations conditioning their synthesis. If we accept this theory we must assume that glucose can also be oxidized without being phosphorylated and that this oxidation can catalyse the synthesis of ATP, which in its turn will promote glucose phosphorylation. Otherwise, the increase in liver ATP observed by the authors after administration of glucose cannot be explained. Indeed, the experiments by Colowick, Kalckar and Cori (11) seem to support this assumption, for they start with free adenylic acid; the primary process in these experiments must necessarily be the building up of ATP.

On the other hand, there is little to show that similar processes occur "in vivo," where ATP synthesis seems to be correlated with the breakdown of hexose-phosphate rather than of glucose and can occur, as has been clearly pointed out by Needham (49), at two different stages of this breakdown. As is well known, adenylic acid by reacting directly with phosphopyruvic acid, is rephosphorylated and regenerates ATP; but the oxidative formation of phosphopyruvic acid itself, or rather of phosphoglyceric acid, from triose-phosphate, is also coupled with resynthesis of ATP, from ADP and free phosphate this time. The recent findings of Warburg and Christian (63) and of Negelein and Brömel (50) (51), explain this coupling, showing that phosphoglyceraldehyde takes up one more phosphate group and that the diphosphoglyceric acid, formed by oxidation of diphosphoglyceraldehyd, subsequently abandons one phosphate group to ADP, converting it to ATP.

An interesting fact emerges from these considerations, namely, that hexose-diphosphate breakdown can catalyse the synthesis of more adenosine-phosphate than is consumed by its own formation out of glucose. The results of Kaplan and Greenberg can therefore be as easily explained by assuming that synthesis of ATP is not prior to glucose phosphorylation but is a consequence of it, through being promoted by the further breakdown of hexose-phosphate. The latter explanation must indeed be true for the muscles, for it has been shown by Leipert and Kellersman (1941) (40), that the accumulation of hexose-monophosphate occurring in the muscles after absorption of glucose goes together with a corresponding loss of ATP. The fact that, in the liver, hexose-phosphate and ATP both increase, might be considered as evidence in favour of Kaplan and Greenberg's theory but it could also be interpreted as a sign that hexose-phosphate is much more rapidly consumed in the liver, which is indeed the case. For we know that glucose is much more rapidly utilized by the liver than by the muscles. Even so, it is not possible to discriminate between both theories, for the advocates of the former might argue that this superiority of the liver is precisely due to the fact that liver tissue possesses more ways of building up ATP than muscle tissue.

If we accept the second alternative, that ATP synthesis is only promoted as a result of hexose-phosphate breakdown, we must find another explanation of the coupling of glucose phosphorylation with pyruvic acid oxidation. This can satisfactorily be done by relating the oxidative process to the preliminary activation of glucose. Activation of a molecule consists in the bringing of this molecule to a higher energetic level and must necessarily be endothermic; a coupling of the kind we surmise appears therefore as a necessity. Moreover, if it be assumed that hexokinase or glycomutin work in conjunction with oxidizing systems, their natural fragility is easily explained.

In short, our actual knowledge concerning glucose phosphorylation allows us to distinguish in it 4 different steps:

1. Adsorption of glucose on glycomutin.
2. Activation of the adsorbed glucose by glycomutin, possibly by conversion into fructo-furanose.
3. Phosphorylation of the activated glucose by ATP.
4. Oxidation of various carbohydrate derivates, principally of pyruvic acid, coupled either with ATP synthesis or with glucose activation.

It is remarkable that an action of insulin on each of these four stages has been described.

Let us consider first the action of insulin on the last of these steps. Early experiments by Polonovski and his co-workers have shown a lowering by insulin of the total amount of intermediate carbohydrate derivates present in the blood. In 1940, activation by insulin of pyruvic acid disappearance was observed simultaneously by von Euler and Högberg (28) and, in our laboratory, by Delrue and De Keyser (15). Various biochemical experiments have also furnished corroborative evidence of an intervention of insulin in the oxidation of carbohydrate by-products. On the other hand, insulin has long been suspected of being an activator of specific cellular oxidation, from the fact that its potency is dependent on the number of disulphide bonds present in its molecule. It is well known that

reduced insulin is totally inactive. A parallel was quickly drawn with glutathione and the natural inference was that insulin could act as hydrogen carrier in cellular oxidations.

These various facts appeared to fit in perfectly with the classical theory, in which the essential action of insulin is described as a furthering of glucose oxidation, and they even supplied a plausible explanation of this action.

According to this "*metabolic theory*," which has been advocated by Polonovski and by many other authors, the main action of insulin consists in an activation of intermediate carbohydrate oxidations and the intensified disappearance of glucose is a consequence of the metabolic "vacuum" thus created.

There are two objections to a theory of that kind. In the first place, even though there is no doubt that insulin has a furthering effect on the oxidation of intermediate products of carbohydrate metabolism, it is clear that this general activity must be related to a highly specific object. Insulin can not be replaced by any other substance, though there are many which are known to activate intermediate oxidations. Glutathione, for instance, does not act the same way as insulin; on the contrary it seems to inhibit insulin action (Levine *et al.* (1939)) (41), possibly because its disulphide-groups take the place of those of insulin and further similar processes without, however, co-ordinating them with a specific system, as insulin does. It must be remembered also that integrity of the disulphide-groups is not the only condition of insulin activity. We must conclude therefore that the action on intermediate metabolism is only one aspect of insulin action; though it is essential in many ways, it is not specific in itself but appears rather like a means through which insulin exerts its specific activity.

In the second place, if we assume the furthering action on oxidations to be the primary action of insulin, then we must admit that insulin promotes blood-sugar disappearance by creating a negative gradient of glucose. But, in that case, there should be a decrease not an increase in hexose-phosphate after insulin and the similarity between the effects of insulin and those of hyperglycaemia becomes difficult to explain. This similarity as well as other experiments which will be mentioned below show that it is the other way about and that insulin creates a positive gradient of glucose. When insulin is given, glucose does not disappear more rapidly because room is made for it, but because more can be stuffed into the same room. Of course during prolonged action, further accumulation of the primary products of glucose absorption will be facilitated by the decongestioning effect of the metabolic action.

Now that we know that oxidation of byproducts of carbohydrate metabolism is a necessary adjunct of glucose phosphorylation, all these difficulties vanish and the specific action of insulin is readily identified. What characterizes insulin and differentiates it from other unspecific activators is that it has the additional property of directing the energy it helps to set free, towards a strictly specific end: glucose phosphorylation. Whether it does so by furthering ATP synthesis or by promoting glucose activation remains to be seen.

To discriminate between these two alternatives, we can apply the same reasoning as before. If the primary action of insulin is to promote oxidative synthesis

of ATP, then it must further glucose absorption by promoting indirectly the phosphorylative disappearance of activated glucose and thus making extra room for more glucose to occupy. On the other hand, if insulin promotes the activation of glucose itself, then it must increase the amount of activated glucose present in the tissues. Experimental evidence seems more in favour of the latter alternative.

It will be remembered that quite a number of early experimental results have been alleged in support of a theory commonly known as the "*permeability theory.*" According to this theory, which rests mainly on results obtained by Loewi and other authors on blood-cells and blood-vessel tissue, the action of insulin consists essentially in a furthering of the adsorption of glucose on the surface of the enzymatic structure responsible for its ulterior metabolism. In another theory also advocated by some authors on the faith of experimental evidence, insulin is said to favour the conversion of ordinary, stable dextro-rotatory glucose into a labile, more levo-rotatory form; in other words, to promote glucose activation. It is true that both these theories have been severely criticized and that the evidence on which they are based has been shown to be inconclusive. It must be noted however that opposite changes have never been asserted to occur and that later experimental results have since been added to their credit.

For instance, a series of experiments by Soula and his co-workers show that, when glucose is administrated, part of it is stored, *as glucose*, in the peripheral tissues, and is subsequently returned to the blood-stream; insulin promotes this storage. According to the authors this storage takes place in the intercellular fluid contained in the lacunar spaces (stockage lacunaire); but this phenomenon may be due as well to a reversible process of adsorption-activation. More recent experiments by Lundsgaard (45), however, do not confirm this storage, but they show that the acceleration of the transfer of glucose from the blood into the muscles cannot be due to a lowering of the glucose concentration inside the muscle-cell, and must therefore be due to an active surface process. The extensive work carried out by Thomas (61) (62) on the rotatory power of blood sugar must also be mentioned in support of the activation theory.

Consequently, the only theory which seems to reconcile all experimental results satisfactorily, is the one in which insulin is considered as an activator of the coupled processes of glucose activation and intermediate carbohydrate derivates oxidation. Indeed, if our first conclusion be accepted that insulin must necessarily act on the same reaction as does the blood-sugar level itself, in other terms on the initial reaction of glucose metabolism:

$$\text{Glucose} \rightarrow \text{X}$$

then we must necessarily accept the biochemical interpretation propounded, which is nothing but the logical inference of the first, as X turns out to be nothing else but activated glucose.

We may therefore conclude that the main action of insulin appears to be an activation of a complex enzymatic system, for which we may retain the name "glycomutin." This system fixates blood glucopyranose and converts it to a

labile, less dextro-rotatory form, possibly fructo-furanose; it borrows the energy necessary to this conversion from the oxidation of various carbohydrate derivates, particularly of pyruvic acid. The activated glucose subsequently reacts with adenosine-triphosphate and is transformed into hexose-monophosphate.

The hexose-monophosphate formed may be either polymerised into glycogen or broken down into pyruvic acid. This breakdown is facilitated by the rupture of the chemical balance between hexose-monophosphate and pyruvic acid, as the former's concentration increases while the latter's initially decreases. Moreover, hexose-monophosphate breakdown is *itself coupled* with ATP synthesis. It follows that the amount of glucose which will be absorbed depends on the velocity of the subsequent breakdown of hexose-monophosphate, which favours it both by removing hexose-phosphate and by regenerating ATP. This breakdown appears to be much slower in the muscles than in the liver. Finally, the pyruvic acid produced from hexose-phosphate will be mainly oxidized, this process involving a corresponding decrease in fat and protein breakdown, but will also be partly used, when an excess of glucose is absorbed, for the synthesis of fat and possibly protein. It will be noticed that insulin, by acting as it does, simultaneously favours the removal of glucose from the blood-stream and the privileged oxidation of its final derivates, while the maximum velocity of glucose absorption under its influence depends essentially on the velocity with which hexose-monophosphate can be converted into those derivates through a chain of reactions which is not, apparently, influenced by insulin.

It is possible that direct oxidation of activated glucose as well as oxidation of other derivates may also play a part in the synthesis of ATP. In that case, a furthering action of insulin on these reactions might possibly also occur.

i. *The action of insulin on protein metabolism.* As has been shown above, all the actions of insulin appear to be satisfactorily interpreted by the theory just described. There exists however one action of insulin which seems to be independent of the others; it is an inhibition of protein hydrolysis, which has been demonstrated by the experiments of Lacquet, De Nayer and Bouckaert (36) (37).

Comparison of their results with those published more recently by Mirsky (1938) (48) enforces a conclusion somewhat different from that originally drawn by the authors. In a first set of experiments, carried out on normal and on pancreatectomized dogs, they have confirmed the fact that the concentration of amino-acids in the blood is raised by pancreatectomy and decreased by insulin. In a second set of experiments, they have compared the blood-amino acid curve of dogs treated with insulin with that of untreated animals, during continuous infusion of various amounts of glycine, and have observed that the mean blood-amino acid level was lowered by insulin in the animals receiving small amounts of glycine, but was raised by insulin in the animals receiving moderately large and large amounts of this amino-acid.

These results tended to show that, as long as the concentration of amino-acids in the blood was sufficiently low to allow a certain degree of proteolysis to proceed in the tissues, addition of insulin inhibited this process, but that, when the blood amino-acid level became high enough to reverse the process and induce proteo-

synthesis in the tissues, the latter process was likewise inhibited by insulin. In other words, insulin appeared to inhibit both the hydrolysis and the synthesis of proteins.

Bouckaert, De Nayer and Cassiman (1934) (6) subsequently showed that, even in the animals receiving small amounts of glycine, the blood amino-acid level was raised by insulin, when the blood-sugar level was maintained by an additional infusion of glucose. These results served to emphasize a second action of insulin on protein metabolism, really dependent on the disappearance of glucose, namely, a strong inhibition of deamination in the liver; as a result of this second action, the removal of amino-acids from the blood-stream is slowed down to such an extent that their concentration in the blood increases despite the simultaneous inhibition of peripheral proteolysis.

The results obtained by Mirsky (1938) (48) confirm the existence of two independent actions of insulin on protein metabolism, one of them being an inhibition of deamination in the liver, the other being an inhibition of proteolysis in the peripheral tissues. But some of the results he has obtained on eviscerated animals are in complete disagreement with the conclusion drawn by Lacquet and his co-workers that insulin inhibits proteo-synthesis as well, for they show that insulin actually promotes the removal of amino-acids from the blood-stream into the peripheral tissues. If we assume that insulin exerts a similar action on the peripheral tissues of the intact animal, another explanation must be found for the fact that the concentration of amino-acids in the blood of the insulinized animal is nevertheless higher than in the non-insulinized animal, when large amounts of glycine are provided. One explanation appears to be that the infusion of additional quantities of glycine affects the hepatic metabolism of the insulinized animal in more or less the same way as the infusion of additional quantities of glucose, and enhances the inhibitory effect of insulin on hepatic deamination. It could be that only a small part of the glycine provided is deaminated and that the furthering by insulin of the oxidation of its non-protein products is sufficient to account for a subsequent inhibition of the deamination of the remaining part. Another possible explanation is that insulin inhibits the removal of amino-acids from the blood-stream by the kidneys, for Mirsky's experiments are carried out on nephrectomized animals.

Whatever the explanation may be, Mirsky's results appear to be quite conclusive and it must be accepted that insulin not only inhibits protein breakdown in the peripheral tissues but promotes the building up of protein from amino-acids as well. This view is strongly substantiated by many recent experiments carried out by Young, Gaebler and other authors, in which it is shown that the sparing and building up of protein, which occurs under the influence of anterior pituitary extracts, requires the presence of sufficient amounts of insulin.

It is to be noted that the inhibition of hepatic deamination by insulin appears to be due largely to the disappearance of extra glucose in the liver, whereas the furthering of protein synthesis in the peripheral tissues seems to be the result of a direct action of the pancreatic hormone on these tissues. The former action may, therefore, be taken as a consequence of the general action of insulin on

glucose metabolism described in the preceding section; the latter appears to be entirely independent of it.

j. *The antagonism between insulin and adrenalin.* It has been shown by De Nayer (18) that adequate quantities of adrenalin may inhibit completely the hypoglycaemic effect of even strong doses of insulin. More recently, de Duve, Marin and Bouckaert (23) have observed that addition of very small doses of adrenalin to the glucose supplied by continuous infusion to insulinized animals leads to a considerable decrease of the quantity of glucose needed to maintain the blood-sugar level normal. In the latter paper, an attempt has been made to reconcile the recently gained information concerning the biochemical mechanism of adrenalin action with that obtained on insulin, in a way which would account for the antagonism between the two hormones. Two facts seem to be of importance:

1. It is remarkable that both insulin and adrenalin increase the amount of hexose-monophosphate present in the tissues. It is not unusual to have two antagonistic factors acting synergically on the formation of one particular substance; the same occurs with the synthesis of hepatic glycogen which is favoured both by insulin and by anterior pituitary extracts or cortin. Indeed part of their antagonism may be precisely due to this apparent synergy, for both the origin of the substance formed and the mechanism of its formation are different when the effect is provoked by insulin or by its antagonist. In the case of hepatic glycogen, we have seen that insulin favours its formation from glucose, by promoting glucose phosphorylation, while anterior pituitary extracts and cortin further its synthesis from protein and possibly fat, by activating glyconeogenesis. In the case of hexose-phosphate, insulin promotes its formation by facilitating the phosphorylation of glucose by ATP, while adrenalin does the same, by activating phosphorolysis of glycogen at the expense of inorganic phosphate. It is evident that activation of one of these mechanisms will necessarily tend to counteract the other.

2. Another important point is that phosphorylation is coupled with oxidative reactions, while phosphorolysis is favoured by reducing processes. Now, we know that insulin also favours the coupled oxidations, particularly that of pyruvic acid. On the other hand, it is well known that adrenalin furthers the formation of lactic acid, which is the product of pyruvic acid reduction; moreover, after a detailed study of the action of adrenalin on glycogen-phosphorolysis, Lee and Richter (39) come to the conclusion that only two hypotheses could account for its mechanism, one of which being that adrenalin furthers the coupling between reducing substances and phosphorylase. We have there a second remarkable instance of the antagonism between the two hormones and a common substrate of their conflicting influences could be cozymase which, in the reduced state, reduces pyruvic acid into lactic acid, and in the oxidized state, intervenes in pyruvic acid oxidation. Recent results by Göbell (31) (32) seem to confirm this assumption, for they show that administration of nicotinamide, one of the main constituents of cozymase, enhances the actions both of insulin and adrenalin.

Conclusion. Except for its independent action on protein synthesis, all the known effects of insulin, as well as the great majority of experimental results obtained on insulin, can be satisfactorily explained by the theory of insulin action put forward in this paper. This theory has been elaborately described (see sections g) and h) and need not be repeated here.

The main points we wish to emphasize are the importance of the blood sugar level, the importance of the hepatic action of insulin, the close similarity between the effects of insulin and those of an increase in blood-sugar level, and the importance of phosphocatalytic metabolism in the metabolism of glucose.

From these various facts, the following picture can be drawn of the part played by insulin in blood-sugar regulation.

The fasting blood-sugar level is the result of an equilibrium between glucose production and glucose consumption. If we assume all neuro-hormonal influences to be stabilized and thus to act in a constant manner, this equilibrium will be essentially a chemical equilibrium, and the resulting blood-sugar level will be the concentration consistent with this equilibrium. The state of equilibrium will be attained when the quantity excreted by the liver will become equal to the quantity consumed by the other tissues. Such a state can always be attained and will automatically be attained by a readjustment of the blood-sugar level, for we know that the total quantity of blood-glucose either consumed or produced is a direct function of the blood-sugar level.

Insulin has the power of modifying this function so that a progressive variation of the amount of insulin present in the tissues will create an infinite number of states of equilibrium, each being characterized by a particular concentration of glucose.

Therefore the basal function of the pancreatic islets will be to maintain a constant level of insulin in the tissues, corresponding to that particular state of equilibrium characterized by a physiological value for the blood-sugar level.

The main antagonists of insulin in this respect are the anterior pituitary hormones and cortin, which act mainly by increasing the production of glucose in the liver, by inhibiting peripheral oxidation of glucose and even by directly opposing the action of insulin.

The second, intermittent action of insulin is to accelerate the return to a normal blood-sugar level whenever hyperglycaemia has occurred, whether through digestive absorption or for any other reason. In that case, an excess of insulin is provided by the pancreas and creates temporarily a new state of equilibrium, characterized by a hypoglycaemic concentration of glucose. The return to a normal blood sugar level will then be accelerated but may be followed by hypoglycaemia. This will be dealt with by a discharge of adrenalin which will reverse the process, and so on.

The main antagonist of insulin, in this urgency function, is, of course, adrenalin.

SUMMARY

In this paper the whole bulk of experimental work carried out on the problem of insulin action at the physiological laboratory of the university of Louvain has

been summarized and discussed in connection with recent data obtained elsewhere. The main points, which will be found summarized under the heading "General theory of insulin action" (p. 58), lead to the conclusion that insulin acts on the chain of reactions which leads, in the liver as well as in other tissues, from blood-glucose to tissue hexose-monophosphate.

The next section is devoted to a close discussion of the biochemical mechanism of this action and it is shown that the only explanation which seems compatible with all the actually known facts is that insulin facilitates simultaneously the adsorption of glucose to a specific enzymatic system responsible for the conversion of glucose into an active, labile form, which is subsequently phosphorylated by ATP, and for the oxidation of various carbohydrate derivates, particularly of pyruvic acid, this process being coupled with the former.

Some aspects of the action of insulin on protein metabolism and of the antagonism between insulin and adrenalin are finally discussed.

We are indebted to Sir Henry Dale, Prof. J. H. Burn and Prof. F. G. Young for their valuable advice and criticism.

REFERENCES

(1) Blalock, A. and M. F. Mason. Am. J. Physiol. **117**: 328, 1937.
(2) Bodo, R and H. P. Marks. J. Physiol. **65**: 48, 1928.
(3) Bouckaert, J. P. and W. Stricker. C. R. Soc. Biol. **91**: 104, 1924.
(4) Bouckaert, J. P., P. P. De Nayer and R. Krekels. C. R. Soc. Biol. **101**: 511, 1929.
(5) Bouckaert, J. P., P. P. De Nayer and R. Krekels. Arch. Int. Physiol. **31**: 180, 1929.
(6) Bouckaert, J. P., P. P. De Nayer and W. Cassiman. C. R. Soc. Biol. **117**: 257, 1934.
(7) Bouckaert, J. P. and P. P. De Nayer. *Invloed van insuline op glucose.* Verh. kon. Vl. Akad. Geneesk. v. Belg., 1946 (in preparation).
(8) Bridge, E. M. Bull. Johns Hopkins Hosp. **62**: 408, 1938.
(9) Britton, S. W. and E. L. Corey. Am. J. Physiol. **131**: 790, 1941.
(10) Cherry, I. S. and L. A. Crandall, Jr. Am. J. Physiol. **125**: 41, 1939.
(11) Colowick, S. P., H. M. Kalckar and C. F. Cori. J. Biol. Chem. **137**: 343, 1941.
(12) Corey, E. L. and S. W. Britton. Am. J. Physiol. **131**: 783, 1941.
(13) Cori, C. F. and G. T. Cori. Proc. Soc. Exper. Biol. and Med. **25**: 66, 1927.
(14) Crandall, L. A., Jr. and I. S. Cherry. Am. J. Physiol. **125**: 658, 1939.
(15) Delrue, G. and J. De Keyzer. C. R. Soc. Biol. **133**: 709, 1940.
(16) De Maeyer, E. and R. De Graer. Act. Biol. Belg. (in press).
(17) De Nayer, P. P. Arch. Int. Pharm. et Thérap. **42**: 461, 1932.
(18) De Nayer, P. P. C. R. Soc. Biol. **111**: 1049, 1932.
(19) Duve, C. de and J. P. Bouckaert. Arch. Int. Pharm. et Thérap. **69**: 485, 1944.
(20) Duve, C. de, P. P. De Nayer, M. Van Oostveldt and J. P. Bouckaert. Arch. Int. Pharm. et Thérap. **70**: 78, 1945.
(21) Duve, C. de, P. P. De Nayer, J. De Keyzer and J. P. Bouckaert. Arch. Int. Pharm. et Thérap. **70**: 383, 1945.
(22) Duve, C. de, H. G. Hers and J. P. Bouckaert. Arch. Int. Pharm. et Thérap. **72**: 105, 1946.
(23) Duve, C. de, G. Marin and J. P. Bouckaert. Arch. Int. Pharm. et Thérap. **72**: 120, 1946.
(24) Duve, C. de, H. G. Hers and J. P. Bouckaert. Arch. Int. Pharm. et Therap. **72**: 129, 1946.
(25) Duve, C. de and H. G. Hers. Acta Biol. Belg. (in press).

(26) DUVE, C. DE AND H. G. HERS. C. R. Soc. Biol. (in press).
(27) DUVE, C. DE. *Glucose, Insuline et Diabète.* Goemaere, ed. (Brussels), Masson et Co (Paris), 1945.
(28) EULER, H. VON AND B. HOGBERG. Naturwissenschaften, 29, 1940.
(29) FIESSINGER, N., H. BENARD, M. HERBAIN, L. DERMER AND G. BAREILLER. C. R. Soc. Biol. 124: 952, 1937.
(30) FISHER, R. A. *Statistical methods for research workers.* Oliver and Boyd (London), 1932.
(31) GOBELL, O. Klin. Wchnschr. 2: 710, 1940.
(32) GOBELL, O. Ber. Phys-med. Ges. Würzburg N. F. 63: 86, 1940.
(33) KAPLAN, N. O. AND D. M. GREENBERG. J. Biol. Chem. 156: 525, 1944.
(34) KAPLAN, N. O. AND D. M. GREENBERG. J. Biol. Chem. 156: 543, 1944.
(35) KAPLAN, N. O. AND D. M. GREENBERG. J. Biol. Chem. 156: 559, 1944.
(36) LACQUET, A., P. P. DE NAYER AND J. P. BOUCKAERT. C. R. Soc. Biol. 115: 434, 1934.
(37) LACQUET, A., P. P. DE NAYER AND J. P. BOUCKAERT. Arch. Int. Pharm. et Thérap. 47: 318, 1934.
(38) LAMERS, K. L. E. C. R. Soc. Biol. 94: 795; 95: 251, 1926.
(39) LEE, M AND D. RICHTER. Biochem. J. 34: 551, 1940
(40) LEIPERT, T. AND E. KELLERSMANN. Hoppe-Seyler's Ztschr. 276: 233, 1942.
(41) LEVINE, R., O. HECHTER, A. GROSSMAN AND S. SOSKIN. Proc. Soc. Exper. Biol and Med. 40: 525, 1939.
(42) LOUBATIERES, A. L. Bull. Soc. Chim. Biol. (zône Sud) 25: 1404, 1943.
(43) LUNDSGAARD, E., N. A. NIELSEN AND S. L. ORSKOV. Skand. Arch. Physiol. 81: 11, 1939.
(44) LUNDSGAARD, E., N. A. NIELSEN AND S. L. ORSKOV. Skand. Arch. Physiol. 81: 20, 1939
(45) LUNDSGAARD, E. Upsala Läk, Förh., N. F. 45: 143, 1939.
(46) MACLEOD, J. J. R. *The control of carbohydrate metabolism.* Bull. Johns Hopkins Hosp. 54: 79, 1934.
(47) MARKS, H. P. AND F. G. YOUNG. J. Physiol. 93: 61, 1938.
(48) MIRSKY, I. A. Am. J. Physiol. 124: 569, 1938.
(49) NEEDHAM, D. M. *Energy yielding reactions in muscle contraction.* Enzymologia 5: 158, 1938.
(50) NEGELEIN, E. AND H. BROMEL. Biochem. Ztschr 301: 135, 1939.
(51) NEGELEIN, E. AND H. BROMEL. Biochem. Ztschr 303: 132, 1939.
(52) NIELSEN, N. A. Skand. Arch. Physiol. 66: 19, 1933.
(53) OCHOA, S. J. Biol. Chem. 141: 245, 1941.
(54) PANY, J. Hoppe-Seyler's Ztschr 272: 273, 1942.
(55) PARNAS, J. K. *Ueber die enzymatischen Phosphorylierungen in der alkoholischen Gärung und in der Muskelglykogenolyse.* Enzymologia, 5: 166, 1938.
(56) REID, C. J. Physiol. 87: 121, 1936.
(57) SOSKIN, S., M. D. ALLWEISS AND D. J. COHN. Am. J. Physiol. 109: 155, 1934.
(58) SOSKIN, S. AND I. A. MIRSKY. Am. J. Physiol. 112: 649, 1935.
(59) SOSKIN, S., H. E. ESSEX, J. F. HERRICK AND F. C. MANN. Am. J. Physiol. 124: 558, 1938.
(60) SOSKIN, S. AND R. LEVINE. Am. J. Physiol. 120: 761, 1937.
(61) THOMAS, J. Bull. Soc. Chim. Biol. 13: 1223, 1931.
(62) THOMAS, J. C. R. Soc. Biol. 112: 1258, 1933.
(63) WARBURG, O. AND W. CHRISTIAN. Biochem. Ztschr 303: 40, 1939.
(64) WIERZUCHOWSKI, M. J. Physiol. 87: 311, 1936.

OPTIMAL GROWTH OF THE RAT[1]

MAX S. DUNN, EDWARD A. MURPHY AND LOUIS B. ROCKLAND

Chemical Laboratory, University of California, Los Angeles

The importance of the rat in studies of growth and nutrition has been universally recognized yet the data on growth and physiological performance have not been reviewed comprehensively since 1924, the date of Donaldson's (3) revised classical treatise. The need for correlation of old and new data is emphasized by the statement of Donaldson (1924, p. x) that the performance of his rats, although a reasonable standard for that period, should not be considered the ultimate criterion. He deduced that "these rats show the characters commonly found in (this and) other laboratory colonies, but they represent neither the best animals nor those in the ideal condition, nor do they necessarily yield the values which will be found in Albinos ten years hence." Abundant evidence in confirmation of this assertion has been provided by the discovery, during the past two decades, of the essential rôle of vitamins, amino acids and other dietary factors in nutrition.

The need for standardization of the performance of rats, stressed by Donaldson and his co-workers at the Wistar Institute, has been appreciated by later investigators, but not all studies of growth, body functions and the nutritional value of food constituents have had fundamental significance. Too often, growth data have been of minimum value because diets were deficient in types or proportions of nutrients. It has been pointed out by Wilson (4) and Needham (5) that, because of the quality of the experimental data, the attempts to validate proposed theories of growth have not proved particularly fruitful.

It has been emphasized by Zucker *et al.* (6) that "probably a majority of the formulations of growth for rats have been based upon or fitted to the data of Donaldson, whose weight-time curves were obtained from rats on a diet poorly characterized and certainly deficient in a number of factors." Zucker and Zucker (7) have stated that "most growth data in the literature which have been used by students of growth probably fail to meet good nutritional standards. The whole growth problem has been confused and more or less run into a blank wall by repeated working over of the same old data. There is a fundamental flaw in these data. New data are needed which do meet reasonable nutritional standards, and a useful, if perhaps limited, start can be made with a single species—the rat."

The inadequacy of the earlier standards of growth was appreciated by Mendel and Cannon (8) who stated in 1927 that the improved growth observed by Osborne and Mendel "implies that, owing to the shortcomings of the rations used, the inherent capacity of the rat to grow has in the past rarely been given full

[1] The writers' work has been aided by grants from the Nutrition Foundation, Inc., and the University of California. Early investigations of the growth process were summarized by Brody (1) in 1927. Laboratory studies of the rat prior to 1939 have been reviewed by Friedman (2).

play in the laboratory. Consequently, the published records and compilations of 'norms' fail to furnish an adequate idea of the rate of growth of which the rat is capable."

Although growth of any type or magnitude may serve to determine that under the observed conditions, a given food constituent functions in some capacity essential to the welfare of the organisms, it has become increasingly apparent that it is far more important to determine, if possible, the types and proportions of food elements required for the maintenance of the highest type of physiological performance. This view that optimal performance of animals is the only truly satisfactory standard for use in measuring the biological value of foods appears to be widespread (9–13).

For present purposes, the terms "optimal" and "normal" are considered to be synonymous and in contrast to "average" or "ordinary."[2] The question, "What is normal or optimal growth or what should be the course of growth in the rat strain employed when there are no nutritional deficiencies?" has been raised by Zucker *et al.* (6, 17) who have stated that "If effects of deficiencies are to be elucidated an unhampered norm should first be worked out. We cannot expect growth on deficient diets simply to reproduce normal growth on a smaller scale." Howe (18) has expressed the similar view that "the kind and rate of feeding has such marked influence on the physical development of animals that it is necessary to define normal growth and physique before attempting to discuss the modifying effects of nutrition."

Normal or optimal growth has been defined (14, 15, 16, 19) as that physiological state in which the animal functions in accordance with its design, actual achievements fulfill all the potentialities of the organism and the growth curve is a pure expression of the inherent growth characteristics. Carlson (13) has defined an optimal diet as "that kind and quantity of food which permits and promotes optimum growth, optimum performance of all biologic functions, optimum resistance to disease, optimum conservation of the factors of safety and powers of repair and optimum length of life with optimum efficiency."

Although growth and the physiological sufficiency of diets are measured inadequately by increase in weight (20), or any other single criterion, it appears probable from Thompson's (21) studies of phylogenetic changes and Huxley's (22) investigations of ontogenetic growth that growth is controlled by a recognizable system of forces. The pattern of growth has not been determined accurately but it has been found that growth data obtained in different laboratories for different strains of rats may be correlated conveniently by means of Huxley's equation.[3] The latter has been reinforced and augmented by the investigations of Zucker and co-workers (6, 7, 17, 23).

[2] Some writers (14–16) have inferred that the terms 'optical," "normal," "average," and "usual" have equivalent meaning.

[3] A discussion of the older growth equations which have been proposed has been omitted since numerous critical reviews (1, 5, 7, 21, 23, 24–33) are available.

In Zucker and co-workers' useful post-weaning growth equation, $\log W = -k/t + \log A$, k and A are positive constants, t is age in weeks from birth and W is weight in grams. A plot

It has been assumed by some investigators that optimal growth and normal growth may not be synonymous terms. Osborne and Mendel (40) pointed out in 1926 that rapid growth may be indicative of distorted development of organs and body parts and that accelerated growth would be physiologically disadvantageous if time alone were gained at the cost of unorderly, and hence detrimental, development. This problem has been studied for more than three decades in the laboratories at Yale University and the Connecticut Agricultural Experiment Station by Osborne and Mendel and their associates including Anderson, Cannon, Hubbell, Moment, Outhouse, Pickens, Smith and Winters.

The present authors' experiments have been predicated on the assumption that the diet and the rats, described in 1932 by Anderson and Smith (41) as the culmination of the nutrition experiments conducted in Osborne and Mendel's laboratories, were near optimal. It is recognized that, at present, the terms "optimal diet" and "optimal physiological state" have only relative significance and there is no intention to imply that other combinations of dietary components may not be of equivalent, or even superior, quality.

There has been no attempt in this review to record all of the voluminous

---

of log W vs 1/t gives a straight line of which k is the slope and, when $1/t = 0$, $\log A = \log W$. The parameters, k and A, permit an analysis of growth in terms of a "growth intensity factor" and an "inherent size factor," respectively. Richards (34) has emphasized Gray's (35) warning that care must be used in interpreting data plotted logarithmically since the diameter of the point on the graph may cover more range than the error of measurement of the variable. It appears that the logarithmic curves of Zucker *et al.* (7) were not subject to this error since it was shown that the relative probable error (ratio of the probable error to the mean) was constant during the whole period of growth.

In Huxley's (22) equation, $\log y = K \log x + \log b$, y is the size or weight of the part, x is the size or weight of the whole, and K (slope of the curve) and b (intercept on the y axis when $x = 1$) are constants. The equivalent expression, $\log y = k_2/k_1 \log x + (\log A_2 - k_2/k_1 \log A_1)$, may be derived from Zucker's equation if the growth of the whole organism is expressed as $\log x = -k_1/t + \log A_1$ and the growth of the part by $\log y = -k_2/t + \log A_2$. The K in Huxley's equation and the ratio k_2/k_1 in Zucker's equation signify that the rate of growth of the part remains constant in respect to the rate of growth of the whole.

Although the constant b is of little biological significance according to Huxley (22) and Needham (36), Zucker *et al.* (7) and Bernstein (33) have pointed out that b is the parameter expressing the ratio of the limiting size of the part to that of the whole, i.e., $b = A_2/A_1^K$. The factors K and b may be useful indices for the characterization and differentiation of species and strains since they are believed to be expressions of the genetic structure and heredity of the organism. Applying Huxley's equation to a large number of vertebrates, Quiring (37) found in 1941 that the K values indicated a fundamental principle of size increase while the b values denoted a scale of being in numerical terms. It might be expected that different strains within the same species would exhibit similar, although smaller, differences.

Zucker *et al.* consider rat growth to be separable into two phases, pre-weaning and post-weaning. For biological reasons, they place the normal post-natal weaning age at 28 days. The heterogonic equation $\log y = k_2/k_1 \log x + (\log C_2 - k_2/k_1 \log C_1)$ was proposed by Huxley (22) and Hamilton and Dewar (38) and it may be derived from Zucker's equation, $\log W = k \log T + \log c$, for the pre-weaning phase of growth or from the related equation of McDowell *et al.* (25). Huxley considered this expression to be in agreement with the work and conclusions of Schmalhausen (39).

literature on the growth of rats on different diets. Consideration has been given, primarily, to the data of the post-weaning period considered to be most highly significant.

TOTAL BODY GROWTH. The data plotted according to Zucker's method in figure 1-A were taken from the papers of Mendel and Hubbell (42) and Anderson and Smith (41). The curves represent the different types of growth obtained in the New Haven laboratories during the period, 1912–1935, with male rats of the same (Yale albino) strain maintained on different diets under essentially the same environmental conditions. The increased rate of growth on the latter diets was accompanied by increased uniformity of growth and by improvement in physiological condition. It was the view of Osborne and Mendel (40) and of Mendel and Hubbell (42) that these effects were due to dietary changes rather than to selective breeding. The superior growth-promoting capacity of the Anderson-Smith diet has been confirmed by Pickens, Anderson and Smith (43). It has been found, also, by Freudenberger (44) that improved diets and care increased the uniformity of growth of rats.

The growth curves obtained by the present authors with the Long-Evans strain of rats maintained on the Anderson-Smith diet are shown in figure 1-B (male rats) and figure 2 (female rats). Other growth curves are given in figure 1 for male rats of the Long-Evans and Wistar strains and in figure 2 for female rats of the Long-Evans strain.

According to Zucker's (45) premises, curves A-3, A-4 and C-2 of figure 1 represent growth on deficient diets. The marked difference between curves C-2 and the two others may indicate, according to Zucker's (17) concept of spontaneous realimentation, that the diets employed differed in respect to the type of nutritional deficiency. The single-straight-line curves A-2, B-2, and C-1 are of the types characterized by Zucker *et al.* (17) as good, normal, optimal and adequate since the growth data obtained with several albino colonies in a number of laboratories yielded straight-line curves of similar slope. Although it was recognized that growth data obtained with diets known to be nutritionally deficient did not yield straight-line curves and that diets which did yield straight-line curves had no pronounced deficiencies, Zucker *et al.* have been careful not to define normal (or optimal) growth as that adhering to their equation. It is of interest that Freudenberger's (46) diet, which yielded straight-line curves (B-2 and C-1 of fig. 1), was later (47) found to be deficient in iodine although the deficiency, as evidenced by rectification, was not sufficient to affect the curve of total body growth.

A type of growth not considered by Zucker and co-workers is represented by curves A-1 and B-1 of figure 1. The difference in slope between the curves A-1 and A-2 for the Yale-albino strain and between the curves B-1 and B-2 for the Long-Evans strain possibly may be explained by the nutritional history of the parent rats. The parents of Anderson and Smith's (41) rats were raised on Mendel and Hubbell's (42) sub-optimal diet which yielded curve A-2 of figure 1. At birth of the young the mothers were placed on the Anderson-Smith diet and at weaning the young were continued on this diet. The young in-

creased rapidly in weight and attained an adult weight optimal for rats of this strain. Similar results were obtained by Pickens, Anderson and Smith (43) under the same conditions. Although their rats at 4 weeks of age were slightly lighter than those of Anderson and Smith, the same adult (16 weeks, post natal) weights were attained by the two groups of rats. On the other hand, curve B-1 represents the growth of rats (maintained on the Anderson-Smith diet) whose

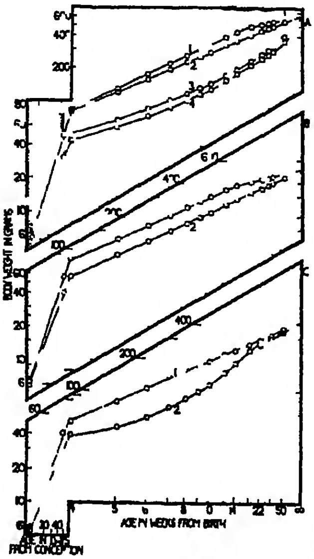


Fig. 1

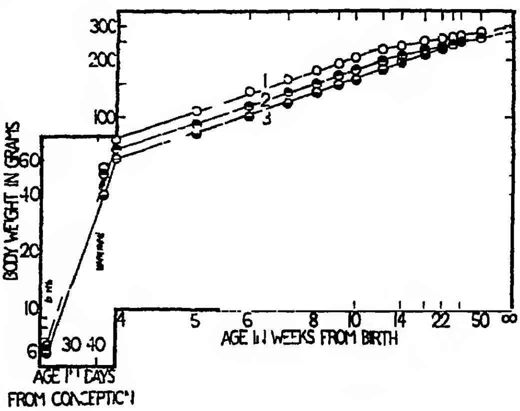


Fig. 2

Fig. 1. Curves showing the relation between body weight and reciprocal age of male rats on various diets. The notations are as follows:

Yale strain. A-1, Anderson and Smith (41), 1932. A-2, Mendel and Hubbell (42), 1935. A-3, Mendel and Hubbell (42), 1925. A-4, Mendel and Hubbell (42), 1910.

Long-Evans strain. B-1, This paper (unpublished data). B-2, Freudenberger (46), 1932.

Wistar strain. C-1, Freudenberger (46), 1932; Greenman and Duhring (60), 1930. C-2, Greenman and Duhring (60), 1923.

Fig. 2. Curves showing the relation between body weight and reciprocal age of female rats on various diets. The notations are as follows:

Long-Evans strain. 1, This paper (unpublished data). 2, Emerson and Evans (56), 1944. 3, Freudenberger (46), 1932.

parents had been raised on the Anderson-Smith diet. If the rats employed by Anderson and Smith had been the progeny of parents which had been raised on the Anderson-Smith diet, it appears probable that a growth curve resembling that found in the authors' laboratory, might have resulted. Hanson and Hayes (48) reported that the birth and weaning weights influenced the adult weight but Anderson and Smith (41) and Zucker *et al.* disagreed with this conclusion. Other workers have found that the adult weight of mice (49), rabbits (50) and Daphnia (51) is not dependent upon the initial growth rate.

If it is assumed that curves A-1 and B-1 of figure 1 represent near-optimal growth it would follow that the single straight-line growth curves are indicative of some dietary deficiency. This hypothesis would seem to be in harmony with the statements of Zucker and Zucker (7) that "the only experimental conditions which can possibly lead to growth according to any relatively simple law require a diet which is adequate for the most rapid growth attainable by nutritional means. Under these conditions only can the growth curve be a pure expression of the inherent growth characteristic of the organism." Since this (or any other) interpretation of the characteristics of optimal growth involve the problem of obesity, the relation of fat deposition to increase in body weight is considered at this point.

It may be inferred that the relatively rapid increase in pre-weaning body weight of rats (curves A-1 and B-1 of fig. 1) raised on the Anderson-Smith or other diet containing a relatively high percentage (20 to 30 per cent) of fat occurs without deposition of fat in excess of the nutritional needs of these animals. Although male rats raised on the Anderson-Smith diet contained a higher percentage of body fat than rats fed the modified (Mendel-Hubbell) Anderson-Smith diet, the data of Pickens *et al.* (43) indicate that the shapes of the body-weight and the fat-free body-weight curves yielded by the two diets were closely similar. The same percentage of body fat is approached as growth on these diets proceeds through maturity. Data in harmony with this observation have been reported by Zucker *et al.* (23) who found that the body weight of rats fed an adequate diet increased slowly up to 70 weeks of age with no evidence (even at 100 weeks of age) of excessive fat deposition, and by Reed *et al.* (52) who observed that the relative distribution of fat in the animal body was the same on high carbohydrate-low fat as on low carbohydrate-high fat diets.[4] The findings of Evans *et al.* (53), Maynard and Rasmussen (54) and Vinson and Cerecedo (55) that lactation is improved on diets containing relatively high proportions of fat may explain the superior quality of such diets for pre-weaning growth.

The straight-line growth curves of the albino colonies considered by Zucker *et al.* were similar in slope. This observation led to the supposition that the slope of the curves provided a mathematical characterization of certain genetical factors, particularly that called "growth intensity." It may be noted from figure 1, however, that the straight-line-portions of different curves may have the same, or different slopes depending, apparently, upon the age and the dietary history of the animals. Zucker *et al.* found less close agreement between the slopes of the growth curves for Long-Evans hybrid rat colonies than for albino rat colonies. These differences were attributed to chance genetical variations resulting from the breeding of hybrid parents but Freudenberger (46) has stated

[4] It has been reported recently by Williams *et al.* (110) that the type of diet [(a) fat, 67.9 per cent of calories; carbohydrate, 9.3 per cent, (b) fat, 38.6 per cent of calories; carbohydrate, 38.6 per cent, (c) fat, 9.3 per cent of calories; carbohydrate, 67.9 per cent] had little, or no, effect upon the proportion of essential lipids to total lipids. The precentage composition of the essential lipids was nearly the same on diets a and b. The percentage composition of the essential lipids observed on diet c differed from that found on diets a and b.

that even the albinos of the Long-Evans strain resemble the Long-Evans strain as a whole rather than the Wistar albinos from which they were derived. This latter view appears to be supported by the fairly close agreement found for the constants of the equations for body-weight increase (curves shown in fig. 2 and those given by Zucker *et al.*) of the Long-Evans rats investigated by Freudenberger (46), Emerson and Evans (56), Zucker *et al.* (7, 17) and the present authors.

The values 640, 540 and 430 for the males and 380, 300 and 280 for the females found for the limiting weights of the Yale-albino, the Long-Evans and the Wistar strains, respectively, denote fundamental size differences of these strains of rats. In general, the constants, calculated by the present authors from the data reviewed by Zucker and Zucker (7), agree fairly well with the values reported by these investigators. These results would not seem to be in agreement with the observation of Vinson and Cerecedo (55) that "an adult female of the Long-Evans strain weighs less than an adult Wistar female."

The logarithmic-growth curves obtained with rats, both of the Yale-albino and the Long-Evans hooded strains, fed on the Anderson-Smith diet changed slope at about 14 weeks of age (no data were available for the females of the Yale strain). There appears to be some factual basis for the inference that changes in the constants of the growth equation at about 14 weeks of age constitute an objective expression of conditions existing at the beginning of a "normally" occurring adult growth period. It is known that rats become physically mature at about 16 weeks and are usually bred at this age (57) and that voluntary activity increases up to the age of 12 to 18 weeks and then declines (58). An analogous growth equation, which changes slope at birth, applies to the prenatal and the post-natal phases of pre-weaning growth (conclusions based on the data of Stotsenberg (59) and Greenman and Duhring (60)).

A different type of superior growth was observed by Evans (61) with rats which had received injections of hypophyseal extracts. Although the increased growth rate of rats, due to dietary improvement alone, reported by Mendel and Cannon (8) and Anderson and Smith (41) compared favorably with that found by Evans and by Bryan and Gaiser (62) with rats injected with anterior pituitary extract, it appears that the inherent-size differences of the rat strains employed were not taken into consideration. The average of the two log weight-reciprocal time curves given by Zucker and Zucker (7) in representing Evans' data for the 1923–24 series of experiments, is shown in figure 3. Even though the curve may be expressed by Zucker's equation it appears that the pattern of growth has been markedly changed by the injection of pituitary extract. The resulting large animals probably were monstrosities since the weights attained were larger than the near-optimal rats of the same strain and the inherently-larger rats of the Yale-albino strain. Other distortions of the growth pattern resulting from the injection of anterior pituitary growth hormone include relatively greater growth of the viscera and glandular organs (63, 64), and of the liver (63, 65, 66). The occurrence of splanchnomicria in pituitary dwarfism and atrophy of visceral organs after hypophysectomy has also been observed (67).

HETERAUXESIS. It has been generally recognised (18, 22, 68–70) that the tissues and organs of the body develop in definite order and proportion relative to the growth of the body as a whole and that they develop at a rate which, to a considerable extent, is determined by the nutritional state. Many investigators believe that abundant nutrition promotes the maximum expression of inherited characteristics while under-nutrition results in retarded development. It seems apparent, however, that many studies on heterogonic growth are subject to the criticism that sub-optimal diets, utilized purposefully in some cases, were employed. Such data may be of value, however, since they may serve to characterize sub-optimal growth.

Skeletal growth, growth of some of the principal organs and changes in body composition with age are discussed in the following sections.

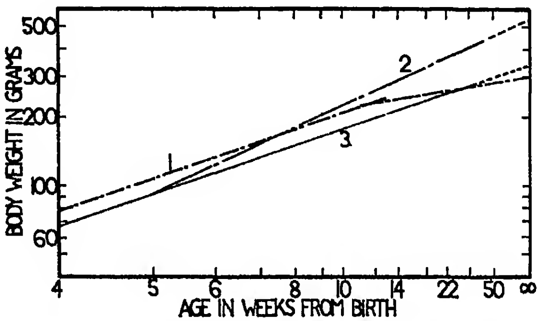


Fig. 3. Curves showing the relation between body weight and reciprocal age of female rats injected with hypophyseal extract. The notations are as follows:

Long-Evans strain. 1, This paper (unpublished data, Anderson-Smith diet). 2, Evans (61), 1923–24. Rats injected with hypophyseal extract. 3, Evans (61), 1923–24. Control rats.

1. *Skeletal growth.* The skeletal growth of male rats of the Yale strain, raised on the diet employed by Osborne and Mendel (40) in 1926, was investigated by Outhouse and Mendel (71) in 1932–33. "Rapid" growth of about 4 grams per day and slow growth of about 2 grams per day were maintained by varying the supplements of yeast and lettuce. Although these investigators found that the general body proportions were nearly the same for both groups of rats, they concluded that body length and weight were more closely correlated than body length and age. At a given body weight the leg bones of the slower-growing (older) rats were longer and chemically more mature than those of the faster-growing (younger) animals. These differences diminished with increasing age and the leg bones of 420 gram rats, both fast- and slow-growing, had the same linear dimensions.

Outhouse and Mendel, as well as Moment (72) who found a similar growth response on the same diets, believed that the growth of their rats was normal on both diets. It would appear, however, from the log-reciprocal plots of Outhouse and Mendel's data (fig. 4) that both diets were sub-optimal, although the diet fed to the faster-growing group (fig. 4-A-1) was less deficient than that supplied to the

slower-growing group (fig. 4-A-2). It may be noted from the log-reciprocal curves shown in figure 4-C and 4-D that the body-length (axial skeleton) and femur-length (appendicular skeleton) curves approach more closely a straight line for rats of the Long-Evans strain fed on the Anderson-Smith diet than for rats of the Long-Evans and the Yale strains fed on other diets.

The influence of age upon the growth of the leg bones is shown graphically by the curves (fig. 5) relating body length and femur length in millimeters. The femurs of Outhouse and Mendel's rats were longer for the slower-growing (older) than for

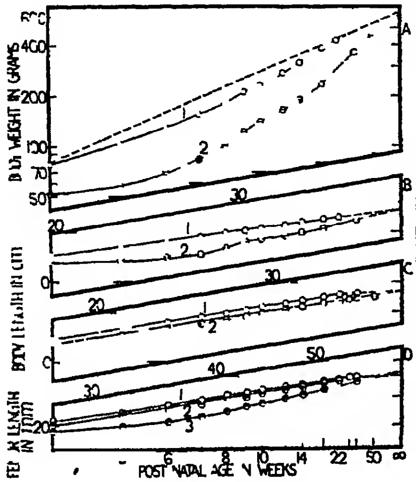


Fig. 4

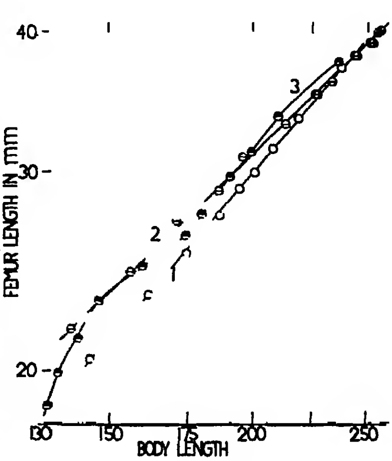


Fig. 5

Fig. 4. Curves showing the relation between reciprocal age and body weight, body length and femur length of male rats on various diets. The notations are as follows:

Yale strain. A-1, B-1 and D-2, Outhouse and Mendel (71), 1932–33. Rapid group. A-2, B-2, and D-3, Outhouse and Mendel (71), 1932–33. Slow group.

Long-Evans strain. C-1 and D-1, This paper. C-2, Freudenberger (46), 1932.

Fig. 5. Curves showing the relation between the femur length and the body length of male rats on various diets. The notations are as follows:

Yale strain. 1, Outhouse and Mendel (71), 1932–33. Rapid group. 3, Outhouse and Mendel (71), 1932–33. Slow group.

Long-Evans strain. 2, This paper.

the faster-growing (younger) group of the same body length. Similar (tachyauxetic) growth curves of the femurs of Wistar-albino rats result from plots of the data obtained by Hammett (73) on Greenman and Duhring's (60) deficient diets. The tachyauxesis of the femur growth observed for rats maintained on deficient diets is in harmony with the earlier report of Winters *et al.* (69) that "while the (deficient) animals made only 8.7 to 19.5 per cent of the normally expected gain in body length, the leg bones made 36 to 41 per cent." No analogous data are available for rats of the Yale and the Wistar strains raised on the Anderson-Smith diet but the authors' data for rats of the Long-Evans strain are shown in figure 5-2. It may be noted that growth of the femur relative to that of the axial skeleton was both uniform and isauxetic ($K = 1$).

It may be concluded from the lateral diameters, the anterior-posterior diameters and the epiphyseal widths of the femurs found by Outhouse and Mendel for Yale rats and by the authors (unpublished) for Long-Evans rats that the thickness of the femurs at a given femur length was greater for the higher- than for the lower-quality diets. Many investigators have reported that the skeleton grew persistently (43, 68–70), but variably (69), on deficient diets. That the curves (fig. 4) of body length and femur length approach more closely to a straight line than those of body weight, is considered to be evidence in harmony with this observation.

It may be concluded that the skeletons of rats raised on the Anderson-Smith diet grow without apparent abnormalities.

2. *Organ growth.* In Moment's (72) investigation of the growth of the principal organs and glands, the strain of rats (Yale) was the same as that employed by Outhouse and Mendel in studying skeletal growth. The diets were similar to those utilized by Outhouse and Mendel. Moment concluded that the size of three different muscles, the kidney, the spleen, the thyroid and the pituitary was more closely correlated with body size than with age. The same correlation was observed for the heart and liver although there was a tendency, at the same body weight, for the rapid-growing group to have larger hearts and livers than the slow-growing group. The growth of the eyeball seemed to be correlated with age rather than body size. At a given body weight, the size of the thymus was greater for the rapid-growing than the the slow-growing rats although the age at which maximum size was attained and began to diminish was the same for both groups.

The size of various organs and glands of the Long-Evans and the Wistar rats at birth, and at 3, 12 and 52 weeks of age was investigated by Freudenberger (46) while data for most of the organs and glands of Wistar rats, based largely on the work of Hatai (74) and Jackson (75), were tabulated by Donaldson (3).

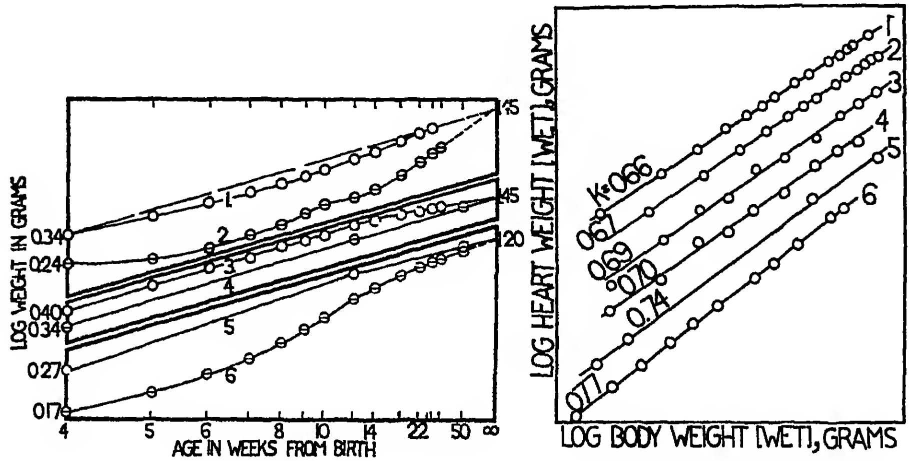
(a) *Heart.* It may be observed from the log heart weight-reciprocal age curves shown in figure 6-A that, according to Zucker's premises, the growth of the heart was sub-optimal for the Yale and the Wistar strains of rats fed on the (presumably inadequate) diets of Moment and Donaldson. It is of interest that growth of the heart appeared to be near-optimal for the Wistar and the Long-Evans strains of rats maintained on the (presumably) adequate or near-optimal diets of Freudenberger and the authors. Even though a reasonably large number of animals was employed, the data reported by Freudenberger for Wistar rats sacrificed for organ measurements and the data reported for his Wistar colony rats were not entirely consistent. For this reason, it is considered that the change in slope of the curve for Freudenberger's Wistar rats may be an artifact.

Log heart weight–log body weight curves are shown in figure 6-B. The points for each strain of rat and for both types (adequate and inadequate) of diets fall reasonably close to a straight line. The lack of any significant difference between the two curves of each pair of curves for the Long-Evans, the Yale and Wistar strains may indicate that, relative to total body growth, the growth of the heart (a vital organ not directly concerned with the metabolism of food) was not measurably influenced by the different diets employed by the investigators

whose experimental data have been recorded. There is some indication, however, of a strain difference in the growth of the heart relative to increase in body weight since the average slopes (K) of the curves were 0.66, 0.70 and 0.76 for the Long-Evans, the Yale and the Wistar strains, respectively.

It appears that Moment's view that, for a given body weight, the heart tended to be slightly larger for a rat which grew rapidly than for one which grew less rapidly, is not substantiated by the data discussed in this section.

It may be concluded that the heart of a rat raised on the Anderson-Smith diet grows without apparent abnormality.



A B

Fig. 6-A. Curves showing the relation between the heart weight and reciprocal age of male rats on various diets. The notations are as follows:

Yale strain. 1, Moment (72), 1933. Rapid group. 2, Moment (72), 1933. Slow group.
Long-Evans strain. 3, This paper. 4, Freudenberger (46), 1932.
Wistar strain. 5, Freudenberger (46), 1932. 6, Donaldson (3), 1924.

Fig. 6-B. Curves showing the relation between the heart weight and the body weight of male rats on various diets. The notations are as follows:

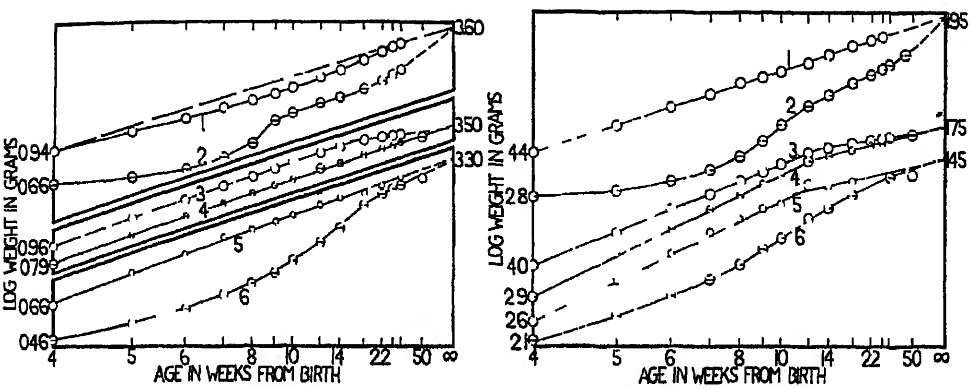
Long-Evans strain. 1, Freudenberger (46), 1932. 2, This paper.
Yale strain. 3, Moment (72), 1933. Rapid group. 4, Moment (72), 1933. Slow group.
Wistar strain. 5, Freudenberger (46), 1932. 6, Donaldson (3), 1924.

(b) *Kidney and liver.* The log kidney weight-reciprocal age curves shown in figure 7-A resemble closely the comparable curves representing heart growth given in figure 6-A. It appears that growth of the kidney was sub-optimal on the inadequate diets and that growth of the kidney on the adequate or near-optimum diets conforms satisfactorily to Zucker's long-weight-reciprocal time equation. It is evident, also, from an inspection of the log kidney weight-log body weight curves (not shown) constructed from these data that kidney growth was more uniform on the higher-than on the lower-quality diets.

It may be observed that the similar liver-growth curves (curves 3, 4 and 5 of fig. 7-B) derived from adequate and near-optimal diets change abruptly in slope

at about 12 weeks of age. This change is considered to be real and not artifactitious since it has been observed not only for liver weight (fig. 7-B) but also for kidney weight (fig. 7-A), body length (fig. 4-C-1), femur length (fig. 4-D-1), and total body weight of female (fig. 2-1 and fig. 2-2) and male (fig. 1-A-1, and fig. 1-B-1) rats.

It is of further interest that the total body and the organs (heart, kidney, and liver) considered as well as the femur and the body have limiting weights or lengths which are characteristic of each of the three strains of rats investigated. These data are given in table 1. It should be emphasized that each figure in the



A B

Fig. 7-A. Curves showing the relation between kidney weight and reciprocal age of male rats on various diets. The notations are as follows:

Yale strain. 1, Moment (72), 1933. Rapid group. 2, Moment (72), 1933. Slow group.
Long-Evans strain. 3, This paper. 4, Freudenberger (46), 1932.
Wistar strain. 5, Freudenberger (46), 1932. 6, Donaldson (3), 1924.

Fig. 7-B. Curves showing the relation between liver weight and reciprocal age of male rats on various diets. The notations are as follows:

Yale strain. 1, Moment (72), 1933. Rapid group. 2, Moment (72), 1933. Slow group.
Long-Evans strain. 3, This paper. 4, Freudenberger (46), 1932.
Wistar strain. 5, Freudenberger (46), 1932. 6, Donaldson (3), 1924.

table was derived from two independent groups of rats of the same strain maintained on different diets in the same or different laboratories.

3. *Body composition.* There has been considerable interest during the past two decades in the growth (increase) of chemical constituents of the animal body with increasing age. Needham (5, 36) has reviewed and correlated a large amount of data on many types of animals and he has proposed what has been called the "chemical ground plan of animal growth." While Needham was concerned, primarily, with embryonic growth, the present discussion deals only with the post-weaning phase of the growth of the rat. Chanutin (76) concluded in 1931 that "chemical maturity" of the rat is reached approximately at the age of weaning (28–30 days) although Moulton (77) had estimated, nearly ten years earlier, that rats become chemically mature at about 50 days of age. The major-

ity of the literature data have been on the chemical groups, protein, fat and ash, although water, chemical elements, creatine, glycogen and other entities have been investigated. The authors' present interest is limited to protein, fat, ash and water.

Earlier investigations (43, 78–81) of the change in chemical composition with age have led to the view that the gain in weight of young growing animals results in large part from accretions of protein, water and ash. It has been observed, as age and live weight increase, that the proportions of protein and water in the animal tend to decrease and the proportion of fat to increase. The proportion of minerals may decrease slightly but does not change greatly under usual conditions. On the other hand, the absolute amounts of all of these substances probably increase with age especially in the case of the rat which grows continuously, although slowly, throughout maturity.

TABLE 1

Limiting size of body and body parts of male rats of different strains

STRAIN	BODY WEIGHT		BODY LENGTH		FEMUR LENGTH*		HEART WEIGHT		KIDNEY WEIGHT		LIVER WEIGHT	
	gm.	*S.R.*	*gm.*	*S.R.*	*gm.*	*S.R.*	*gm.*	*S.R.*	*gm.*	*S.R.*	*gm.*	*S.R.*
Yale albino	640		27.5		44		1.45		3.60		19.5	
		7.5		3.7		5.0		0		1.2		3.6
Long Evans hooded	540		26.5		41		1.45		3.50		17.5	
		6.2		6.3		3.3		7.3		4.3		5.6
Wistar albino	430		23.5		39		1.20		3.30		14.5	

*Air-dried. All other values wet weight.

S.R. is the significance ratio (difference divided by the probable error of the difference). If the quotient is greater than 3 the difference is considered significant.

The effect of the plane of nutrition on body composition is illustrated by McMeeken and Hammond's investigations of swine, reviewed by Howe (18), and by Moulton, Trowbridge and Haigh's (80) studies of cattle in which different nutritional levels were maintained by controlling the quantity of food ingested. Similar data for the body composition of rats have been reported by Pickens, Anderson and Smith (43) who maintained different nutritional levels by controlling the quality of diets fed *ad libitum*. Essentially the same relations of protein, fat, ash and water were found for both species of animals and for both types of nutritional control.

Moulton (77) expressed the opinion that the plane of nutrition would not affect the composition of animals calculated on a fat-free basis unless the conditions were such as to cause retardation of development. A similar view was advanced by Pickens, Anderson and Smith who found that the whole bodies varied considerably in composition at 230 days of age but that the fat-free material of all groups was practically of identical composition. Different degrees of correspondence were observed, however, over the experimental period from weaning to 230 days. It has been found from semi-logarithmic plots of the

TABLE 2

Percentage of body components of male rats on various diets

WATER			PROTEIN			ASH			FAT			BODY PORTION	DIET	RAT STRAIN	REFERENCE
Days†			Days			Days			Days						
37	90	127	37	90	127	37	90	127	37	90	127				
73	70		21	24		4.4	5.1		1.7	0.9		Vital	‡	Wistar	Nash (82)
67	66	65	15	15	16*	3.0	3.8	3.8	12	12	13	Body	Gliadin	Yale	Pickens *et al.* (43)
69	67	66	16	18	18	2.9	3.8	3.8	7.9	8.0	9.6	Body	Lactalbumin	Yale	Pickens *et al.* (43)
67	63	61	16	18	19	2.4	3.0	3.2	12	13	14	Body	Stock	Yale	Pickens *et al.* (43)
65	60	57	16	16	16	2.2	2.5	2.6	14	19	22	Body	Anderson-Smith	Yale	Pickens *et al.* (43)
		70			21			2.7			4.9	Body	Low-salt	Yale	Light *et al.* (82)
		64			20			3.9			11	Body	Calorie control	Yale	Light *et al.* (82)
68		62	17		19	2.4		3.4	12		15	Body	Normal control	Yale	Light *et al.* (82)
60	63		15	17		1.5	1.3		20	18		Reducible	‡	Wistar	Nash (83)

* Changed to lactalbumin diet at 110 days of age.
† Postnatal age.
‡ Described as "supermaintenance diet," composition not given.

TABLE 3

Percentage (fat-free) of body components of male rats on various diets

WATER			PROTEIN			ASH			BODY PORTION	DIET	RAT STRAIN	REFERENCE
Days†			Days			Days						
37	90	127	37	90	127	37	90	127				
75	71	70	21	24	24	4.5	5.2	5.2	Vital	‡	Wistar	Nash (83)
76	75	75	17	17	18*	3.4	4.3	4.4	Body	Gliadin	Yale	Pickens *et al.* (43)
75	73	73	18	19	20	3.2	4.1	4.2	Body	Lactalbumin	Yale	Pickens *et al.* (43)
76	72	71	18	21	22	2.7	3.5	3.7	Body	Stock	Yale	Pickens *et al.* (43)
76	73	73	19	20	20	2.6	3.1	3.3	Body	Anderson-Smith	Yale	Pickens *et al.* (43)
		74			22			2.8	Body	Low salt	Yale	Light *et al.* (82)
		72			22			4.4	Body	Calorie control	Yale	Light *et al.* (82)
77		73	19		22	2.7		4.0	Body	Normal control	Yale	Light *et al.* (82)
75	77		19	21		1.9	1.6		Reducible	‡	Wistar	Nash (83)

* Changed to lactalbumin at 110 days of age.
† Postnatal age.
‡ Described as "supermaintenance diet," composition not given.

data of Pickens *et al.* that the relation between the various components on the different diets was the same whether calculated on a total-weight or a fat-free basis. Differences in composition were of reduced magnitude when calculated on the latter basis. Examination of the plots reveals that the rats fed the Anderson-Smith diet had higher fat content of their bodies and more nearly optimal skeletal growth (ash) than rats maintained on deficient diets.

It was concluded by Light, Smith, Smith and Anderson (82) in 1934 that, under unfavorable nutritive conditions, the chemical composition of the body may be varied so that the available food material is best used to support the vital equilibria and to maintain the essential tissues. This view that changes in body composition may reflect an "adaptation to deficient diets" was reiterated in the later studies of Pickens, Anderson and Smith (43). The general concept that there may be a "vital portion" and a "reducible portion" of the animal body has been discussed more recently by Nash (83). This hypothesis is somewhat analogous to that proposed in 1919 by Terroine (84). The vital portion was considered to be the part necessary for life and the reducible portion the part normally in excess of the minimum requirement for existence. There was assumed to be a vital and reducible portion of each chemical constituent equivalent, respectively, to the amount present at death from inanition and the amount lost during the period of inanition leading to death. Nash reported data illustrating the characteristics of the vital and reducible portions of Wistar rats and he concluded that these body portions differed in composition. On the basis of this concept, it may be assumed that the body composition of rats on deficient diets would resemble more closely the composition of the vital than the reducible portion.[5]

Values for the percentage composition of water, protein, ash and fat in the vital and reducible portions of rats calculated from Nash's data and in the bodies of rats calculated from the data of Pickens *et al.* (43) and Light *et al.* (82) are shown in tables 2 and 3. Although both the Yale and Wistar strains of rats were employed in these studies, definite strain differences in respect to chemical composition have not been demonstrated unequivocally even though the possibility that such differences may be found is indicated by the work of other investigators. Strain differences of rats in response to dietary choline (85), biotin (86,87), lactose (88,89), glucose (90–94) and the B-vitamins, minerals and proteins (95) have been reported.

[5] It may be that the method employed by Nash does not give a true picture of changes in ash content of animals maintained on inadequate diets. The decrease in skeletal material in starvation may not be equivalent to the increase on a poor diet although the ash values may approach each other. Nash has stated that "it is conceivable that by varying the inanition diet, e.g., by restricting in the diet only the constituent considered, the reducible portion of the constituent may be found to be larger and the vital portion smaller than these respective portions as determined by the effects of a complete inanition diet, since on a complete inanition diet death may occur due to the insufficiency of an essential chemical other than the particular constituent under consideration." He pointed out, in addition, that death during starvation is not necessarily due to loss of body substance *per se*, but may be caused by inability to maintain physiological equilibria.

TABLE 4

K values (slopes) of heterauxetic curves for body components of male albino rats on various diets

TOTAL-WEIGHT BASIS				FAT-FREE BASIS			BODY PORTION	DIET	RAT STRAIN	REFERENCE
Water	Protein	Ash	Fat	Water	Protein	Ash				
0.97	1.15	1.13	0.63	0.97	1.15	1.11	Vital	*	Wistar	Nash (83)
0.98							Body	Not given	Wistar	Lowrey (105)
0.97							Body	Not given	Not given	Hamilton and Dewar (38)
0.97		1.11	0.75	0.93		1.09	Body	Not given	Wistar	Hatai (109)
0.97	1.15	1.28	1.09	0.97	1.15	1.28	Body	Gliadin	Yale	Pickens *et al.* (43)
		0.88				0.93				
0.93	1.23	1.33	1.43	0.97	1.15	1.40	Body	Calorie control	Yale	Pickens *et al.* (43)
	0.98	0.82				0.90				
0.95	1.11	1.07	1.10	0.97	1.13	1.07	Body	Control	Wistar	Nash (83)
	1.11	1.00					Body	Not given	Not given	Buchner and Peter (106)
		1.43								
0.95	1.11	1.04	1.54	0.97	1.13	0.93	Body	Not given	Not given	Chanutin (76)
						1.26				
0.95	1.11	1.28	1.11	0.95	1.13	1.33	Body	Stock	Yale	Pickens *et al.* (43)
		0.74	1.84			0.87				
0.92	1.07	1.19	1.17	0.95	1.11	1.35	Body	Anderson-Smith	Yale	Pickens *et al.* (43)
			1.80							

* Described as "supermaintenance diet," composition not given.

It may be concluded from the data in table 2 that, as the efficiency of the diet increases, the composition of the body becomes less similar to that of the vital portion and more similar to that of the reducible portion. On a fat-free basis (table 3), the percentage of water and protein is essentially the same for all diets but that of ash is inversely related to the efficiency of the diet. Additional information is yielded by consideration of the curves of growth and heterogony derived from Huxley's heterauxetic equation which has been employed by Needham and other investigators (5, 23, 36, 38, 83, 96). Some of the data examined indicate that on adequate diets the increase of chemical constituents with time follows this equation at least up to maturity. The data are not extensive, however, and this problem has not been studied intensively although there is some indication that "disturbances" occur particularly during transition periods between growth phases (pre-weaning and post-weaning growth periods). The curves obtained by Nash who plotted the growth of the elements of the vital portion against the vital portion as a whole, are of particular interest since nutritionally limited growth should resemble that of the vital portion.

In table 4 are listed values for the slopes of the straight-line curves obtained from plots of the logarithm-substance weight against the logarithm-body weight both on the wet-weight and the fat-free basis. These values are given in the order of decreasing similarity to the slopes calculated by Nash for the water and protein components of the vital portion. In some cases a single straight line is inadequate to represent all of the data. Needham (5) pointed out that heterauxetic curves which are not single straight lines may occur but he excluded curves of this type from his discussion. It has been found in the present review that many of the variant curves may be attributed to nutritional disturbances. The K value for each section of such curves is shown in the table.

It is of interest that of the diets employed by Pickens, Anderson and Smith the Anderson-Smith diet was the only one which yielded a single straight-line heterauxetic curve for body ash and it may be inferred from the data given in table 4 that the proportion of water, protein and ash to total body weight increased with an efficiency which was at least as great on the Anderson-Smith as on the other diets. It may be concluded that rats raised on the Anderson-Smith diet are not abnormal in the proportions of water, protein and ash to total body weight on a fat-free basis.

Physiological state. It has been pointed out, previously, in this paper that the ideal conception of optimal growth would include optimal reproductive ability, optimal resistance to disease and optimal longevity. Although McCay (97) has found that increased rate of growth may be accompanied by decrease in longevity, his observed longevity was not related to optimal resistance to unfavorable environmental conditions. No reports are available on the longevity or the resistence to disease of rats raised on the Anderson-Smith diet.

In table 5 are presented data on the reproductive performance of the Yale, the Long-Evans and the Wistar strains of rats maintained on various diets. It appears evident from these figures that the reproductive performance was best on the Anderson-Smith diet. In view of the findings of Griffith and Farris

(57), Greenman and Duhring (60), Morrison (98), Fitzhugh *et al.* (99), Templeton *et al.* (100) and Emerson and Evans (56), it may be found that reproduction beyond first litters may not be as satisfactory because of the relatively high fat (lard) content of the Anderson-Smith diet. Experiments designed to test this hypothesis are in progress in the authors' laboratory.

Exercise is included among the more important environmental factors which, in addition to diet, may influence growth and physiological well-being of experimental animals. Since all of the rats considered in this review were housed in stationary cages, there is no information available on the extent to which the growth and reproduction data obtained with the Anderson-Smith and other

TABLE 5

Average reproductive performance of different strains of rats on various diets

FEMALES FERTILE	YOUNG					STRAIN	REFERENCE
	Per litter	Birth weight	Weaned†	Weaning weight			
				Males	Females		
per cent	*number*	*gm.*	*per cent*	*gm.*	*gm.*		
86	7.2		71	23	26	Yale	Mendel and Hubbell (42), 1912
65	6.3		67	31	31	Yale	Mendel and Hubbell (42), 1919
68	6.4		76	31	30	Yale	Mendel and Hubbell (42), 1925
93	9.6	5.8	90	48	47	Yale	Mendel and Hubbell (42), 1935
81	5.9	5.8	82	43‡		Long-Evans	Evans and Bishop* (107), 1923
	6.7	5.9		40	39	Long-Evans	Freudenberger (46), 1932
95	9.3	6.3	100	56	54	Long-Evans	This paper*
	6.1	4.2		20	22	Wistar	Donaldson* (3), 1924
81	8.2	4.8	67	29‡		Wistar	Greenwood* (108), 1940
	8.5	5.4		41	38	Wistar	Freudenberger (46), 1932

* First litter data. Litter number not specified in other cases.

† Refers to the 6 or 7 young retained of the total litter.

‡ Average of young weaned. Figures not given for each sex.

diets would have been altered by voluntary exercise. The conclusions on this point from published data are conflicting. It was reported by Hatai (101) that the heart, liver and kidneys of exercised rats were heavier than the same organs of non-exercised animals though there was little, if any, difference in body length and body weight. The spleen and the lungs of the non-exercised rats were found to be heavier than these organs of the exercised animals. Hatai's observation that exercise prevented, or at least delayed, the onset of a prevalent pulmonary infection in the albino rat may be interpreted to mean that the heavier lungs of the unexercised rats resulted from infection rather than from any normal physiological variation. Seven generations of rats were studied by Donaldson (102) who found that exercised males were slightly lighter and exercised females somewhat heavier than the non-exercised animals although

the effect was not cumulative from generation to generation. Body length was not appreciably affected, but the heart, kidney, suprarenals and gonads of the exercised rats were somewhat larger than these organs of the non-exercised animals. Donaldson attributed the decrease in body weight of exercised males to loss of body fat and this conclusion was corroborated by Reed, Yamaguchi, Anderson and Mendel (52). Borovansky (103) reported that the body weight both of male and female rats was heavier for the exercised than the non-exercised animals. Cages facilitating voluntary exercise were recommended by Greenman and Duhring (60) who stated that, "as a result of a highly varied diet, gentling and exercise, it is not only possible to produce heavier animals but their fertility also may be greatly increased." Some evidence indicating decreased longevity of exercised rats has been presented by Slonaker (58).

The effect of food intake, activity and temperature upon the body weight of twelve female albino rats of the Sprague-Dawley strain has been studied recently by Brobeck (104). It was concluded that (a) there was a negative correlation between weight change and activity when food intake and environmental temperature were constant, (b) increasing the food intake increased body weight gain when activity and environmental temperature were constant, and (c) weight loss was less at 86° than at 70°F. when food intake and activity were constant. The following criticisms of these carefully controlled experiments might be offered. Weight gains would not be comparable at different ages (4 to 10 months) of rats since growth on the experimental diet, even of the non-exercised rats, was sub-optimal. It would be of interest, and possibly of increased physiological significance, to carry out this same experiment with a near-optimal ration fed *ad libitum*.

SUMMARY

Rats, rat organs and chemical components of rats grow (increase in weight or other dimensions) at rates or in proportions which are considered to be sub-optimal, adequate or near-optimal. The straight-line curves derived from plots of the logarithms of weights or lengths and the reciprocals of time according to the equations of Zucker *et al.* appear to provide satisfactory, although empirical, bases for the evaluation of growth of rats and the relative efficiency of experimental diets. Rats of the Yale, the Long-Evans and the Wistar strains appear to have widely different limiting weights which represent strain differences necessary to consider in growth experiments. Although voluntary exercise and temperature affect gain in body weight of female rats significantly on sub-optimal diets it has not been determined to what extent growth on near-optimal diets would be influenced by these factors. No abnormalities in the growth of rats, rat organs and chemical components of rats maintained on the Anderson-Smith diet have been observed. It appears that the Anderson-Smith diet is superior to the other diets for which experimental data are available in promoting high quality growth and reproductive performance of rats.

It has been concluded as a working hypothesis that the Anderson-Smith diet contains the types and proportions of chemical substances which are near-

optimal for the growth of body, organs and body components and for the reproduction of the rat. No information is available on the efficacy of this diet in promoting longevity, resistance to disease and physiological well-being as determined by other criteria.

REFERENCES

(1) Brody, S. Missouri Agric. Exper. Sta. Res. Bull., no. 97, 1927.
Brody, S. Bioenergetics and growth. New York, 484–663, 1945.
(2) Friedman, M. H. Yearbook of Agr., U. S. Dept. Agr., 145, 1939.
(3) Donaldson, H. H. Boas Anniversary Volume, New York, 1906.
Donaldson, H. H. The rat. Memoirs Wistar Inst. Anat. Biol., Philadelphia, (1915 and 1924).
(4) Wilson, E. B. Symposia on Quant. Biol. 2: 199, 1934.
(5) Needham, J. Biochemistry and morphogenesis. London, 531, 1942.
(6) Zucker, L., L. Hall, M. Young and T. F. Zucker. Growth 5: 399, 1941.
(7) Zucker, L. and T. F. Zucker. J. Gen. Physiol. 25: 445, 1942.
(8) Mendel, L. B. and H. C. Cannon. J. Biol. Chem. 75: 779, 1927.
(9) Sherman, H. C. The science of nutrition. New York, 7, 1944.
(10) McLester, J. S. Nutrition and diet in health and disease. Philadelphia, 247, 1943.
(11) Bacharach, A. L. and J. C. Drummond. Chem. Ind. 59: 37, 1940.
(12) Reed, O. E. Yearbook of Agr., U. S. Dept. Agr., 1045, 1939.
(13) Carlson, A. J. Science 97: 385, 412, 1943.
(14) King, C. D. Yale J. Biol. Med. 17: 493, 1945.
(15) Marston, W. M., C. D. King and E. Marston. Integrative psychology. New York, 434, 437, 1931.
(16) Gumpert, M. You are younger than you think. New York, 50, 1944.
(17) Zucker, T. F., L. Hall, M. Young and L. Zucker. J. Nutrition 22: 123, 1941.
(18) Howe, P. E. Yearbook of Agr., U. S. Dept. Agr., 469, 1939.
(19) Dunn, M. S. Symposium on the biological value of proteins. Div. Agr. Food Chem., Amer. Chem. Soc., Cleveland, April 6, 1944.
(20) McCollum, E. V. and M. Davis. J. Biol. Chem. 20: 415, 1915.
(21) Thompson, D. W. On growth and form. London, 1943.
(22) Huxley, J. S. Problems of relative growth. New York, 1932.
(23) Zucker, L., L. Hall, M. Young and T. F. Zucker. Growth 5: 415, 1941.
(24) Schmalhausen, J. Arch. f. Entwicklungsmech. 108: 322, 1926.
(25) McDowell, E. C., E. Allen and C. G. McDowell. J. Gen. Physiol. 11: 57, 1927.
(26) Weinbach, A. B. Growth 5: 217, 1941.
(27) Needham, J. Chemical embryology. London, 1931.
(28) Lotka, A. J. Elements of physical biology. Baltimore, 1925.
(29) Snell, G. D. Proc. Nat. Acad. Sci. 15: 274, 1929.
(30) Murray, H. A. J. Gen. Physiol. 1: 39, 1926.
(31) Heilbrunn, L. V. Outline of general physiology. Philadelphia, 229, 1943.
(32) Davenport, C. B. Symposia on Quant. Biol. 2: 203, 1934.
(33) Bernstein, F. Symposia on Quant. Biol. 2: 209, 1934.
(34) Richards, O. W. Papers from Tortugas Lab., Carnegie Inst. Wash. 29: 173, 1936.
(35) Gray, J. Brit. J. Exper. Biol. 6: 248, 1929.
(36) Needham, J. Biol. Rev. 9: 79, 1934.
(37) Quiring, D. P. Growth 5: 301, 1941.
(38) Hamilton, B. and M. M. Dewar. Growth 2: 13, 1938.
(39) Schmalhausen, J. Arch. Entwicklungs-Mech. 109: 455; 110: 33, 1927.
Schmalhausen, J. Biol. Zentralbl. 50: 292, 1930.
(40) Osborne, T. B. and L. B. Mendel. J. Biol. Chem. 69: 661, 1926.
(41) Anderson, W. E. and A. H. Smith. Am. J. Physiol. 100: 511, 1932.

(42) MENDEL, L. B. AND R. B. HUBBELL. J. Nutrition 10: 557, 1935.
(43) PICKENS, M., W. E. ANDERSON AND A. H. SMITH. J. Nutrition 20: 351, 1940.
(44) FREUDENBERGER, C. B. Anat. Rec. 56: 47, 1933.
(45) ZUCKER, T. F. AND L. ZUCKER. Ind. Eng. Chem. 35: 868, 1943.
(46) FREUDENBERGER, C. B. Am. J. Anat. 50: 293, 1932.
(47) FREUDENBERGER, C. B. AND F. W. CLAUSEN. J. Nutrition 15: 1, 1938.
(48) HANSON, F. B. AND F. HAYS. Anat. Rec. 35: 83, 1927.
(49) CROZIER, W. J. AND E. V. ENZMAN. J. Gen. Physiol. 19: 249, 1935.
(50) MERRELL, M. Human Biol. 3: 37, 1931.
(51) ANDERSON, B. G., H. LUMER AND L. J. ZUPANEIC, JR. Biol. Bull. 73: 444, 1937.
(52) REED, L. L., F. YAMAGUCHI, W. E. ANDERSON AND L. B. MENDEL. J. Biol. Chem. 87: 147, 1930.
(53) EVANS, H. M., S. LEPKOWSKY AND E. A. MURPHY. J. Biol. Chem. 106: 431, 1934.
(54) MAYNARD, L. A. AND E. RASMUSSEN. J. Nutrition 23: 385, 1942.
(55) VINSON, L J. AND L. R. CERECEDO. Arch. Biochem. 3: 389, 1944.
(56) EMERSON, G. A. AND H. M. EVANS. J. Nutrition 27: 469, 1944.
(57) GRIFFITH, J. R. AND E. J. FARRIS. The rat in laboratory investigation. Philadelphia, 1942.
(58) SLONAKER, J. R. J. Animal Behavior 2: 20, 1912.
SLONAKER, J. R. J. Comp. Neurol. and Psych. 17: 342, 1907.
(59) STOTSENBURG, J. M. Anat. Rec. 9: 667, 1915.
(60) GREENMAN, M. J. AND F. L. DUHRING. Breeding and care of the albino rat for research purposes. Wistar Inst. Anat. Biol., Philadelphia, 2nd ed., 1931.
(61) EVANS, H. M. Harvey Lect. 19: 212, 1923-24.
(62) BRYAN, A. H. AND D. W. GAISER. Am. J. Physiol. 99: 379, 1932.
(63) PUTMAN, T. J., E. B. BENEDICT AND H. M. TEEL. Arch. Surg. 18: 1708, 1929.
(64) CUSHING, H. AND L. DAVIDOFF. Monographs of The Rockefeller Institute, New York, no. 22, 1927.
(65) LEE, M. O. Relation of the anterior pituitary growth hormone to protein metabolism in the pituitary gland. Baltimore, 193, 1938. Quoted by LEE and FREEMAN (66).
(66) LEE, M. AND W. FREEMAN. Endocrinology 26: 493, 1940.
(67) SMITH, P. E. Am. J. Anat. 45: 205, 1930.
(68) JACKSON, C. M. The effects of inanition and malnutrition upon growth and structure. Philadelphia, 1925.
(69) WINTERS, J. C., A. H. SMITH AND L. B. MENDEL. Am. J. Physiol. 80: 576, 1927.
(70) SMITH, A. H. J. Nutrition 4: 427, 1931.
(71) OUTHOUSE, J. AND L. B. MENDEL. J. Exper. Zool. 64: 357, 1932-33.
(72) MOMENT, G. B. J. Exper. Zool. 65: 359, 1933.
(73) HAMMETT, F. S. J. Biol. Chem. 64: 409, 1925.
(74) HATAI, S. Am. J. Anat. 15: 87, 1913.
(75) JACKSON, C. M. Am. J. Anat. 15: 1, 1913.
(76) CHANUTIN, A. J. Biol. Chem. 93: 31, 1931.
(77) MOULTON, C. R. J. Biol. Chem. 57: 79, 1923.
(78) ARMSBY, H. P. AND C. R. MOULTON. The animal as a converter of matter and energy. New York, 1925.
(79) HANKINS, O. G. AND H. W. TITUS. Agric. Yearbook, U. S. Dept. Agr. 458, 1939.
(80) MOULTON, C. R., P. F. TROWBRIDGE AND L. D. HAIGH. Mo. Agric. Exper. Sta. Res. Bull., no. 55, 1922.
(81) HAECKER, T. L. Univ. Minn. Agric. Exper. Sta. Bull., no. 193, 1920.
(82) LIGHT, U. E., P. K. SMITH, A. H. SMITH AND W. E. ANDERSON. J. Biol. Chem. 107: 689, 1934.
(83) NASH, C. B. Growth 6: 151, 1942.
(84) TERROINE, E. F. Ann. Sci. Nat. Zool. 4: 5, 1920.

(85) COPELAND, D. H. Proc. Soc. Exper. Biol. Med. **57**: 33, 1944.
(86) NIELSEN, E. AND C. A. ELVEHJEM. J. Biol. Chem. **144**: 405, 1942.
(87) KENNEDY, C. AND L. S. PALMER. Arch. Biochem. **7**: 9, 1945.
(88) MITCHELL, H. S. J. Nutrition **12**: 447, 1936.
(89) ERSHOFF, B. H. AND H. J. DEUEL. J. Nutrition **28**: 225, 1944.
(90) COLE, V. V. AND B. K. HARNED. Endocrinology **23**: 318, 1938.
(91) ORTEN, J. M. AND H. B. DEVLIN. J. Biol. Chem. **136**: 461, 1940.
(92) COLE, V. V., B. K. HARNED AND C. E. KEELER. Endocrinology **28**: 25, 1941.
(93) SAYERS, G., M. SAYERS AND J. M. ORTEN. J. Nutrition **26**: 139, 1943.
(94) COWGILL, G. R. Physiol. Rev. **25**: 664, 1945.
(95) NOBACK, C. R. AND H. S. KUPPERMAN. Proc. Soc. Exper. Biol. Med. **57**: 183, 1944.
(96) TEISSIER, G. Trav. Stat. Biol. Roscoff **9**: 29, 1931.
(97) McCAY, C. M. J. Nutrition **18**: 1, 1939.
(98) MORRISON, F. B. Feeds and feeding. 12th ed., Pasadena, California, 149, 1943.
(99) FITZHUGH, O. G., A. A. NELSON AND H. O. CALVERY. Proc. Soc. Exper. Biol. Med. **56**: 129, 1944.
(100) TEMPLETON, G. S., F. G. ASHBROOK AND C. E. KELLOGG. Conservation Bull., no. 25, Fish and Wildlife Service, U. S. Dept. Interior 17, 1942.
(101) HATAI, S. Anat. Rec. **9**: 647, 1915.
(102) DONALDSON, H. H. Am. J. Anat. **50**: 359, 1932.
(103) BOROVANSKY, L. Bull. Intern. Acad. Sci. Boheme, 1930. Quoted by DONALDSON, (102).
(104) BROBECK, J. R. Am. J. Physiol. **143**: 1, 1945.
(105) LOWREY, L. G. Anat. Rec. **7**: 143, 1913.
(106) BUCKNER, G. D. AND A. M. PETER. J. Biol. Chem. **54**: 5, 1922.
(107) EVANS, H. M. AND H. S. BISHOP. J. Met. Res. **3**: 201, 1923.
(108) GREENWOOD, M. L. Thesis, Iowa State College, Ames, 1940.
(109) HATAI, S. Am. J. Anat. **21**: 23, 1917.
(110) WILLIAMS, H. H., H. GALBRAITH, M. KAUCHER AND I. G. MACY. J. Biol. Chem. **161**: 463, 1945.

DEVELOPMENT OF THE GRAAFIAN FOLLICLE AND OVULATION

FREDERICK L. HISAW

The Biological Laboratories, Harvard University

Twenty years have passed since it was discovered that ovulation could be induced in sexually immature rats and mice by making single or repeated implants of pituitary tissue (Smith, 1926; Aschheim, 1926; Zondek, 1926). Also, a similar period has elapsed since a practical laboratory technique was devised for hypophysectomizing rats (Smith, 1927). These two events gave impetus to an enormous amount of research on pituitary-ovarian relationships. Many species of vertebrate animals, ranging from fishes to mammals, have been used in experiments on ovulation and the gonadotropic preparations employed have been equally diverse. Much of this work has been of an empirical nature and further complicated by wide variation both in methods used for preparing and standardizing gonadotropic materials and experimental procedures. This makes it difficult, if not impossible, to find a common ground for the comparison of results.

The object of this review is to discuss the normal processes involved in the development and ovulation of a Graafian follicle with particular emphasis on the parts played by the pituitary and ovarian hormones. For this purpose, a rather arbitrary selection from the extensive literature dealing with this subject was made and only those papers were included that have a direct bearing on the development of the topic. The author is aware that this may lead to errors of omission and consequently, wherever possible, works of general nature are included from which additional references may be obtained.

Follicular development. Ovulation, in mammals, marks the end of the follicular phase of the estrous cycle and the beginning of the luteal phase. When considered from the standpoint of probable evolutionary significance the luteal phase is a specialization. Ovulation is not followed necessarily by the development of a corpus luteum and when such occurs it is not always associated with viviparity. Corpora lutea are not present in amphibia but are found in many species of oviparous, ovoviviparous and viviparous elasmobranch fishes and reptiles. Such corpora are not essential for the maintenance of pregnancy nor is it known that they secrete progesterone. Therefore, neither luteinization nor the presence of progesterone, at least in vertebrates other than mammals, is an indispensable requirement for ovulation.

On the other hand, the factors that control development of the follicle and ovulation seem to be fundamentally the same, if not identical, in all vertebrates. The physiological processes involved in the follicular phase of the estrous cycle of mammals are apparently homologous with those responsible for growth of follicles and ovulation in species of vertebrates in which a luteal phase is lacking. Thus, the available evidence indicates that in this regard all vertebrate animals are basically alike.

Ovarian function in the adult animal is under the direct hormonal control of the anterior lobe of the pituitary gland (Smith, 1939) and the two active agents involved are the follicle stimulating hormone (FSH) and the luteinizing hormone (LH) (Fevold, 1939, 1943; Chow, 1943). The ability of the ovaries to respond to these pituitary gonadotropins is attained as the animal aproaches sexual maturity while the ovaries of the very young animal are refractory. It is also a common observation that even in the adult animal all the follicles are not equally responsive. When a pituitary extract is given in small doses to a rat the effects may be limited almost entirely to the larger follicles, but if the dosage is increased, smaller follicles are called into the reaction, while the primordial follicles fail to be stimulated even when large amounts of the preparation are injected. This raises the question as to the point at which, in the life of a follicle, it becomes responsive to gonadotropic hormones and thus comes under pituitary control. Therefore, some consideration should be given to the conditions governing follicular growth since they determine whether or not development is terminated by a successful ovulation.

One of the striking features of the adult ovary is the presence of tissues that have retained their embryonic characteristics. The cells of the germinal epithelium, in probably all vertebrates, are capable of active mitoses throughout the reproductive life of the animal. This tissue seems to be homologous with the coelomic epithelium and either lies entirely on the surface of the ovary or, as is more usual, is represented by tissue in the ovarian cortex which was derived earlier from the surface epithelium (Mossman, 1938). The embryonic nature of the germinal epithelium is also shown by the fact that oogenesis may continue even after sexual maturity (Hartman, 1939; Mossman, 1937).

The primordial follicles, whether they are budded off directly from the germinal epithelium on the surface of the ovary or are formed in the underlying ovarian cortex, at first do not have a theca interna. Such follicles consist of an oocyte surrounded by a single layer of granulosa cells. When the oocyte begins to grow, a theca interna is formed from the contiguous connective tissue. As the theca thickens by hyperplasia and the addition of more cells from the stroma, it is provided with vascular loops which cover the follicle but do not enter the granulosa. Correlated with this is an increase in mitotic activity in the granulosa and the ovum becomes situated in the center of a solid ball of cells several layers in thickness.

Oogenesis and follicular development in mammals can proceed up to this point in the absence of pituitary hormones. In the immature hypophysectomized rat follicles may develop to the stage of beginning antrum formation (Smith, 1930) and the total number of primordial follicles may be greater than normal (Swezy, 1933). Hypophysectomy of the adult animal leads to rapid atresia of all of the large antra-containing follicles, while the small follicles without antra may show no effects. This seems to be the general situation among mammals but in some species follicular development, in the absence of the pituitary, does not progress as far as it does in the rat (see Smith, 1939, for references).

The absence of pituitary control over oogenesis and early development of the

follicle is probably characteristic of all vertebrates. Smith (1939) found that hypophysectomy of frog larvae (*R. boylei*) did not prevent growth of the gonads up to the stage normally attained at the time of metamorphosis. Similar observations have been made on other species of amphibia (Burns, 1932; Atwell, 1933; and Witschi, 1937) and Shapiro and Shapiro (1934) report not only the degeneration of large ova in toads (*Xenopus laevis*) but also an increase in the number of small follicles.

We know almost nothing in this regard about reptiles and only a few observations have been made on birds. Hill and Parkes (1934) found no follicles of a greater diameter than 1 or 2 mm. in the ovary of a hypophysectomized hen and state that the condition produced is essentially similar to that found in the non-laying bird at the height of the moult. Probably the most important difference found among the vertebrates is the size the ovum can attain after ablation of the hypophysis. The mammalian ovum can reach full size but the growth of large ova in other groups is limited by an inability to deposit yolk.

The sequence of events seen in the formation of primordial follicles in the mammalian ovary is suggestive of developmental processes observed in the organization of tissues in an embryo. The similarity is so striking that one is tempted to postulate that it is probably brought about through the co-ordinated interaction of a self-contained system of organizers within the growing follicle. The actual existence or nature of such organizers in the follicle has not been demonstrated, but organizers are known to be present in the developing gonad (Witschi, 1939). The arrangement of the granulosa of a primordial follicle is apparently under the control of the ovum. When the ova are destroyed in mice and rats by x-ray treatment the germinal epithelium may form cords and clumps of granulosa cells instead of follicles (Parkes, 1926, 1927; Brambell, Fielding and Parkes, 1928; Drips and Ford, 1932; Allen, Hisaw and Gardner, 1939). Although there is much confusion concerning the details of the effects of x-rays on the ovary and the histogenesis of tissues following a treatment, yet it seems that aggregates of granulosa cells can organize the connective tissue of the stroma into structures resembling a theca interna. If this is true, then it is clearly the granulosa and not the ovum that is responsible for the formation of the theca. At the same time the theca interna provides the follicle with a blood supply and either this or a stimulus derived from the theca itself may account for the increase in mitotic activity of the granulosa in a growing follicle.

It has been shown that estrogen has a direct effect on the ovary and this is of additional importance since there is considerable evidence that the theca interna is at least one of the sites of formation of estrogen (Corner, 1938; Mossman, 1937). Bullough (1942) found that intraperitoneal injections of estrogen greatly increased mitotic activity and oogenesis in the germinal epithelium in such distantly related vertebrates as mice and minnows (*Phoxinus laevis*). He suggests that the increased mitosis in the granulosa of a growing follicle might also be due to estrogen. It is interesting that estrogen can act as an organizer at its source of formation, and also function in the co-ordination of more remote bodily processes.

Regardless of the mechanism follicular growth can proceed in mammals, as

has been mentioned, up to the appearance of the antrum folliculi in the absence of the pituitary. This condition is reached in the normal mouse by the twelfth to fourteenth day of age (Engle, 1931) and in rats by the eleventh day (Hargitt, 1930). Previous to this time, the young follicles in both rats and mice are refractory to the gonadotropic action of pituitary hormones. A similar situation exists in rabbits (Hertz and Hisaw, 1934). Probably the ovaries of all mammals are refractory to gonadotropins from birth to a specific age which is characteristic of each species. Clauberg (1932) speaks of this period as "infantile" and the interval between the appearance of responsiveness and the first estrus as "juvenile." The morphological difference between a follicle that will not respond to pituitary hormones and one that has gained competence to react is not obvious. However, it seems that the acquirement of sensitivity is correlated with the time at which the cells of the theca interna normally start differentiating into epithelioid tissue.

Another indication that a follicle has emerged from the "infantile" into the "juvenile" condition, and consequently is susceptible to pituitary stimulation, is the presence of an antrum. The nature of the conditions under which antra are formed is not fully understood but the results of experiments by Lane (1935) are suggestive. Lane first made a statistical study of the ovaries of normal rats from 15 days of age to sexual maturity and of rats hypophysectomized at 28 days of age and killed from one to 38 days later. The larger follicles were classified into two groups; those having two or more layers of granulosa cells were called primary follicles and those having antra were designated as vesicular follicles. These data were used in determining the effects of pituitary hormones and estrogen on the ratio between these two classes of follicles and their total number.

Lane found that 22 day old rats given FSH twice daily for three days and killed when 28 days of age showed an increase over normal in the total number of primary and vesicular follicles, but the percentage of vesicular follicles remained about the same. When LH was used the total follicular number varied little from that of the untreated animal but there was a marked increase in the percentage of those follicles having antra. These gonadotropic preparations also produced similar modifications in the ratio between primary and vesicular follicles in hypophysectomized rats (Lane and Greep, 1935). These results indicate that the formation of follicular antra in rats is associated with the action of LH. It is quite probable that LH plays an important part in this process but whether it alone is directly responsible for producing this effect has not been satisfactorily proven.

It also seems possible that estrogen may influence follicular development by causing an increased secretion or release of LH from the pituitary. Hohlweg (1934) and Hohlweg and Chamarro (1937) found that luteinization could be produced in the ovaries of immature rats by injecting estrogen. Lane (1935) also found that an estrogen treatment caused an increase in the percentage of antra containing follicles, in the ovaries of immature rats, similar to that observed following the administration of LH. Other observations that tend to support the opinion that estrogen stimulates the secretion of LH by the pituitary will be presented in the subsequent discussion of ovulation.

The immediate point under discussion concerns the factor or factors that enable a young follicle to gain competence to respond to the gonadotropic hormones of the pituitary and thus start on its way toward full maturity and ovulation. There are very good reasons for assuming that FSH, LH and estrogen are present at the time the ovary of a young animal gains competence to respond to injected pituitary gonadotropins. Yet we are quite ignorant concerning the individual rôles played by these hormones in the conditioning processes in a young follicle. It is evident that the earlier experiments just cited, though important and suggestive, fall short of the mark in that they do not establish the individual action of a particular hormone to the exclusion of all others. The pituitary gonadotropic preparations were admittedly not pure and that FSH and LH were separated completely is unlikely. Recently FSH and LH have been isolated as pure proteins and should prove to be of great value in the solution of this problem. (For references see Chow, 1943, and Fevold, 1943.)

What seems to be needed are experiments in which the separate and combined actions of these gonadotropins, and estrogen, are tested on the ovaries of hypophysectomized animals. Some very important investigations of this nature have been made but so far more emphasis has been placed on changes in ovarian weights and gross morphology than on the follicular populations and cytological modifications of the ovary.

Greep, van Dyke and Chow (1942) studied the effects of LH (metakentrin, ICSH) and FSH (thylakentrin) isolated as pure proteins from hog pituitaries, on the ovaries of hypophysectomized immature rats. The LH preparation prevented atrophy in the interstitial cells when the treatment was started soon after the operation or would repair the atrophic changes when the treatment was begun several days post-operatively. It also produced enlargement of the thecal cells and at high doses converted some of the smaller follicles into pseudo-lutein bodies. These authors made little point of the presence of antra other than to mention their absence in the healthy-appearing follicles in control animals killed twelve days after hypophysectomy and the presence of microscopic antra in a few follicles in the ovaries of animals that had received LH.

The administration of pure FSH caused the Graafian follicles to grow, mature, and pass into atresia without becoming luteinized or cystic. The ovaries remained small even in animals given large doses and the follicles did not attain the size of mature follicles in normal rats. The interstitial tissue and theca interna were not affected and there was no indication of estrogen secretion. Results that agree with these in most respects have been reported for pure gonadotropins obtained from sheep pituitaries (Fraenkel-Conrat, Li and Simpson, 1943).

When hypophysectomized rats were given a combination of pure FSH and LH simultaneously the effects on the ovaries were like those produced by unfractionated pituitary extracts (Greep, van Dyke and Chow, 1942). The ovaries were heavier than could be obtained with either hormone alone. The follicles grew to full size and were luteinized while the theca interna was well developed and estrogen was secreted. These experiments confirm previous ideas as to the functions of the pituitary gonadotropins in follicular development (Fevold, 1939)

but they leave many questions unanswered and at the same time raise a new one concerning the possible importance of the direct action of estrogen on the follicle.

Williams (1940, 1944, 1945) recently has made the observation that tablets of diethylstilbestrol implanted subcutaneously in immature female rats two days after hypophysectomy prevent or greatly retard the atrophy of the ovaries. He also found that the response of such animals to the injection of pregnant mare serum was markedly increased and the luteinization of the granulosa was much more pronounced than in the nonestrogenized controls. Pencharz (1940) also has reported an increase in size of the ovaries of hypophysectomized rats given subcutaneous implants of diethylstilbestrol. The increase in ovarian weight in response to diethylstilbestrol was about three times that produced by estradiol dipropionate while testosterone propionate was without effect. It was suggested that this discrepancy between the action of the two estrogens probably was due to a difference in rate of absorption.

These estrogens, and especially diethylstilbestrol, had a marked effect on the follicles. The ovaries of the treated animals were packed tightly with medium-sized follicles and the interstitial tissue was greatly reduced. Pencharz also made the interesting observation that the reaction to estrogen gave the otherwise refractory ovaries of the hypophysectomized animal a most remarkable ability to respond to chorionic gonadotropin.

Simpson, Evans, Fraenkel-Conrat and Li (1941) confirmed the work by Pencharz and also studied the effects of estrogens on the response of the ovaries to FSH and LH preparations. They noted that the increase in size of the ovaries of hypophysectomized rats given estrogen was due to a "slight but definite" growth of many follicles. Such follicles in spite of an increase in size did not develop antra and it is mentioned that some showed a peripheral zone of degeneration. This last point was not discussed in detail but it appears that the defect (ref. fig. 4) was in the basal part of the granulosa. The estrogens did not repair the interstitial tissue and presumably this was also true for the theca interna although it was not described.

The influence of diethylstilbestrol on the action of LH and FSH was tested on immature female rats that were hypophysectomized at 28 days of age. Estrogen treatment was started the day after hypophysectomy, the gonadotropins were given on the fifth to seventh days and the animals were killed on the ninth. The estrogen did not modify the action of LH, the reaction being the additive effects of both hormones. However, there was a marked increase in the reaction of the ovaries to FSH. FSH alone produced development of large and medium follicles and there was no repair of the interstitial tissue. When the same dosage of FSH was given to animals that had received diethylstilbestrol the ovaries attained a much greater weight and contained large follicles, corpora lutea and thecal luteinized follicles while the interstitial tissue "was not repaired (or only partly)." The luteinization was probably due to the presence of LH as Fraenkel-Conrat et al. (1943) state that they had not been successful in eliminating all of the luteinizing hormone (ICSH) from their FSH preparations. However,

these results indicate that estrogens greatly potentiate the action of FSH in the hypophysectomized animal and at the same time, judging from dosage, suggest the possibility that the treatment made the theca interna and interstitial tissue more sensitive to the action of LH.

These observations on sexually immature animals suggest several points of interest that can be discussed better after follicular growth in the estrous cycle of the adult has been considered.

In the normal estrous cycle of guinea pigs and rats those follicles that are destined to ovulate begin their development during a period of estrus and reach maturity at the succeeding estrus. During the intervening diestrum they maintain a growth-rate that is nearly constant (Dempsey, 1937; Boling, Blandau, Soderwall and Young, 1941). This probably holds for most mammals in which the estrous periods follow each other in more or less rapid succession and as far as the history of an individual follicle is concerned, after growth has started, this seems true in a general way for all vertebrates. The chief difference being a matter of time.

Some very significant facts regarding follicular growth have been obtained by examining the mitotic rate in the granulosa and theca through the use of colchicine. Schmidt (1942) made a thorough study of the follicles in the ovaries of guinea pigs, from the time of their origin in the germinal epithelium (Schmidt and Hoffman, 1941) to ovulation and the formation of corpora lutea. Primary follicles 25 to 30 μ in diameter are formed from cell nests in which the flattened, undifferentiated granulosa cells become arranged around the young ovum. This process seems to take place without mitosis and mitotic figures are rarely seen in follicles less than 60 μ in diameter, at which time the flattened granulosa cells are differentiating into the typical cuboidal type.

In the guinea pig the follicles acquire a thin theca interna as they approach a diameter of 200 μ and at this stage, or soon afterwards, a small antrum is formed. At this point, as indicated by mitotic activity, the follicle enters a period of active growth, presumably under pituitary influence. At first this is slow but becomes rapid when the follicle has reached a diameter of about 300 μ. The growth-rate in both granulosa and theca is most active from this time until the follicle is 600 μ in diameter. This is followed by a gradual decline in mitotic activity in the theca while in the granulosa the number of mitoses continues to increase, though growth is no longer as rapid, until a follicular diameter of 800 μ is attained. Thus it is seen that the period of rapid mitoses in the granulosa is preceded by the formation of a theca interna and when mitotic activity reaches its peak in the theca and begins to decline this change apparently is reflected in the granulosa by a decrease in mitotic rate.

These peaks of mitotic activity doubtless indicate an abrupt turning point in the physiology of the follicle. Those that escape atresia continue to increase in size, primarily through the accumulation of follicular fluid, but the number of mitoses rapidly decreases and they are absent at ovulation. Follicles having a diameter greater than 800 μ are first seen on the eleventh day of the cycle. This is very significant, as the eleventh day is also the average time that guinea pigs

come into estrus when the cycle is shortened by removing the corpora lutea after ovulation (Loeb, 1911; Dempsey, 1937). Also Dempsey, Hertz and Young (1936) found that ovulation could be induced in the guinea pig as early as the twelfth day of the cycle by injecting an LH preparation. Therefore, the follicle becomes "mature," in the sense that it is capable of ovulation, soon after its diameter increases beyond 800 μ. However, in the normal cycle the sudden spurt in growth, or preovulatory enlargement, may not occur until three or four days later.

These observations on the guinea pig probably represent the general situation in mammals. The only other animal that has been studied with an equal degree of thoroughness is the rat. Lane and Davis (1939) employed the colchicine technique in a careful study of rat ovaries selected at various times during the estrous cycle. Their results differ from those mentioned for the guinea pig only in matters of degree in that the events seen in the sixteen-day cycle of the guinea pig occur more rapidly in the four to six-day cycle of the rat (Boling, Blandau, Soderwall and Young, 1941).

It is obvious that the facts at hand are inadequate as a basis for definite conclusions concerning the physiological mechanism directing and co-ordinating the formation of a follicle, yet they present certain suggestive features that encourage speculation. In the first place it seems probable that the follicle is under two sets of controls, one within the ovary itself and the other outside. The endogenous factors in the ovary work on the principle of embryological organizers. For example, the primordial ovum seems able to arrange the cells that surround it into a granulosa one cell in thickness. The infrequent mitoses in such follicles indicates that the growing ovum probably exercises little if any influence on such activity. However, the fact that during atresia and preovulatory enlargement preceding ovulation, the last place in which mitotic activity disappears is the germ hill, suggests that the mature ovum possesses some ability to stimulate mitoses.

The organizing action of the ovum also is indicated by the observation that after the egg is killed by x-rays, the granulosa cells remain in clumps and cords. Also, the fact that the granulosa can persist in the absence of an ovum makes possible further deductions. Ovaries of x-rayed animals secrete estrogen and if the theca interna is the source of estrogen we may conclude that it is the granulosa rather than the ovum that is responsible for organization of the theca. However, what influence if any progesterone might have on the organization of a follicle in the estrous cycle of mammals is as yet unknown.

The active agents formed outside the ovary are the pituitary gonadotropic hormones FSH and LH which, by their combined action, cause the secretion of estrogen within the ovary. The results of the investigations that have been cited definitely show that from the time a follicle enters the period of rapid growth until ovulation it is constantly under the influence of these three hormones. Furthermore, these observations also give some indication of the function of each hormone. Maintenance and maturation of the theca interna are controlled by LH. Also, FSH cannot promote the secretion of estrogen in the

absence of LH. It seems probable, though it has not been proven, that the theca interna through the action of LH acquires competence to respond to FSH.

One of the functions of FSH is to produce enlargement of the follicle, primarily by the accumulation of fluid in the antrum. This action is limited in the hypophysectomized rat as shown by the failure of the follicles to reach normal size (Greep, van Dyke and Chow, 1942). The reason for this is not clear nor has the source of the follicular fluid been determined. Neither is it known with certainty which structure, if any, in the follicle gives a specific response to FSH. It has been generally thought that the granulosa was the structure concerned but this probably should be questioned now since it is known that hyperplasia of the granulosa can be induced with estrogens and that this effect greatly increases the response to FSH. This, of course, raises the question as to the importance of estrogen, which is always present, during the simultaneous action of FSH and LH.

A point worth considering is the possibility that estrogen secreted by the theca interna in response to FSH and LH may function in follicular development as the growth stimulating agent for the granulosa. This thought is not invalidated by the fact that estrogens must be administered in large doses to produce this effect (Williams, 1944). The theca interna closely invests the granulosa and the constant secretion of minute amounts of estrogen would be many times more effective under such a situation than much larger amounts given subcutaneously. Also, the fact that estrogen is secreted in the ovary may be a good physiological reason for a much higher threshold for ovarian structures, if such exists, than for responsive tissue elsewhere in the body.

Another possibility worth mentioning before concluding this line of thought is that some estrogen may be secreted in the ovaries of very young or hypophysectomized animals. The presence of lipoidal material in the thin theca of small follicles less than 200 μ in diameter, as demonstrated by histochemical techniques, lends support to this idea (Dempsey and Bassett, 1943). Follicles in the rat's ovary do not respond to pituitary gonadotropins before they have a diameter of about 200 μ and Williams finds that the granulosa of primordial follicles is not stimulated by estrogens. It is also true that mitosis in such follicles normally proceeds at a relatively slow rate. These observations make one wonder if the granulosa gains competence for rapid growth in response to estrogens at the same time that the follicle becomes sensitive to gonadotropins. If this is true, then it seems possible that the pituitary gonadotropins, for one thing, simply take over and augment a process that already has been started.

A question that often is raised is why a juvenile animal whose ovaries are capable of being stimulated by injected gonadotropins, does not come into sexual precocity in response to her own pituitary secretion. The histology of the ovaries and the changes that occur following pituitary ablation clearly show that the pituitaries of such animals are actively secreting. It is also true that the ovaries may show several cycles of follicular growth and atresia before the first estrus.

We have no quantitative data that are satisfactory for an analysis of these

facts. Several investigators have mentioned the possibility of an increase in secretion of the gonadotropins as the animal approaches maturity. There are no acceptable techniques for determining the exact amount of FSH and LH in pituitary tissue or unfractionated extracts (Fevold, 1939) and until such are devised this point cannot be settled. Another possibility worth considering is that the follicles may become more responsive to a given amount of pituitary gonadotropins with the advance in age of the animal. It has been shown that this is true for the reactivity of the male and female accessory reproductive organs to estrogens and androgens (Hooker, 1942; Price and Ortiz, 1944; Price, 1944).

We know considerable about the physiology of follicular development but there is much yet to be learned about specific processes involved. One of the obstacles that has retarded advancement has been the lack of pure preparations of pituitary gonadotropins but this now has been removed. Pure LH has been prepared in three laboratories (Chow, van Dyke and Greep, 1942; Li, Simpson and Evans, 1942; Fevold, 1943) and FSH in one (Chow, 1943). It is true that so far the reactions produced in the ovary by these pure substances tend to confirm previous reports for FSH and LH preparations, in which the two hormones were not completely separated. However, pure FSH and LH have not been used in making a critical analysis of follicular development and their effects on the follicular populations of the ovary should be determined. This would contribute to a better understanding of the synergistic reaction produced in the ovary by the action of the two hormones. It has been demonstrated that growth of the granulosa in response to estrogen facilitates the response of the follicle to FSH, but it has not been shown that the changes produced in the theca interna by pure LH will also potentiate FSH. All that is known, in this respect, for the pure gonadotropins is that the response of the ovary is greater when FSH and LH are given simultaneously; however this does not eliminate the participation of estrogen in the reaction. A detailed study of morphological changes in the follicle induced by the independent and concurrent actions of the pure gonadotropins and estrogens would greatly advance our knowledge of ovarian physiology.

Ovulation. The rupture of a mature Graafian follicle and release of an ovum has been the subject of more research during the last twenty years than almost any other problem concerned with ovarian physiology. Out of this have come certain well established facts, one of which is that in normal animals ovulation is under the control of the pituitary. Also, the reaction can be induced by injecting the gonadotropins obtained from the urine of pregnant women and the blood serum of pregnant mares.

This review is concerned with development of the follicle and ovulation in the normal estrous cycle; therefore we are interested in pituitary-ovarian interactions. However, experiments in which the chorionic gonadotropins were employed have contributed much toward a better understanding of normal ovulation, and are useful for our discussion. Yet, one should keep in mind the fact that chorionic gonadotropins are hormones of pregnancy and are neither chem-

ically nor physiologically identical either with each other or with pituitary gonadotropins. Their gonadotropic action, though interesting and useful for experimentation, is nevertheless incidental to their normal function (Engle, 1939; Astwood and Greep, 1939; Hisaw, 1944).

The rupture of a follicle, as seen in mammals, is preceded by a period of rapid growth referred to as pre-ovulatory enlargement. A similar phenomenon probably occurs in other vertebrates but is modified and obscured by the presence of large yolk-laden ova. The morphological changes that take place during pre-ovulatory enlargement and rupture of the follicle have been described in detail for several species. The literature dealing with this has been reviewed by Hartman (1939) and the reader is referred to this and a more recent paper by Dawson (1941).

The follicles that undergo pre-ovulatory enlargement and ovulation in the estrous cycles of guinea pigs and rats are those that have survived a period of extensive atresia (Schmidt, 1942; Lane and Davis, 1939). Such follicles can be distinguished by the third day of the cycle in the rat and by the twelfth day in the guinea pig. From this time until pre-ovulatory enlargement sets in there is a progressive increase in diameter and a further differentiation in the theca and granulosa. The most conspicuous change seems to involve the theca and consists of an increase in blood supply, shown by a rapid proliferation of endothelial cells which becomes enormous during pre-ovulatory enlargement.

The number of eggs ovulated at one time is quite constant for the different species which raises the question as to why certain follicles are selected and others are eliminated by atresia. There is no obvious answer but it seems possible that the immediate cause of atresia might result from a failure of proper differentiation of the theca interna. This thought seems more plausible when one considers the differential action of FSH and LH on the ovary. Fevold (1937) found that the ovaries of hypophysectomized rats showed a greater response to an FSH preparation when the treatment was preceded by three daily injections of LH. Although the LH preparation did not cause an increase in ovarian weight, the effects it produced augmented the response to FSH. The gross appearance of the ovaries indicated that the reaction to FSH was facilitated by an increase in the number of follicles that were capable of responding.

There are a number of facts that tend to support the idea that atresia is due to defective differentiation of the theca interna resulting from inadequate stimulation by LH. The first of course is the specific action of LH on the theca interna resulting in hypertrophy of the thecal cells and increase in blood supply. Another is the theca as a source of estrogen (Corner, 1938) and the part it might take in the growth and differentiation of the granulosa. If estrogen is responsible for growth of the granulosa then the observation that indications of atresia first appear in that structure also suggests failure of the theca (Mossman, 1937; Schmidt, 1942).

The fact that atrophy of the follicles in hypophysectomized rats is prevented by the direct action of injected estrogen on the granulosa also supports the argument (Williams, 1945). The effect of estrogen on the ovary resembles that of

LH in that it apparently increases the number of follicles that are responsive to FSH, yet it differs in a very important respect. Although LH alone prevents atrophy of the theca and promotes its differentiation it cannot maintain the follicles for an extended period, the probable reason being that estrogen is not secreted. It is logical to infer, from experimental evidence, that the well-being of a follicle depends upon the co-ordinated action of FSH, LH, and estrogen. A recent contribution to this thought is the observation by Williams (1940) that estrogen increases the sensitivity of the granulosa to the luteinizing action of pregnant mare serum. Whether or not this explanation for atresia of large follicles previous to ovulation is entirely correct, it seems evident that such follicles do not become atretic necessarily because they are incapable of responding to developmental stimuli.

Superovulation. That many more follicles are capable of being brought to ovulation than normally occurs has been known since the early experiments of Smith and Engle (1927). They obtained 20 to 48 ovulations in rats and mice by implanting pituitary tissue. Such ova were capable of being fertilized and as many as 29 implantations were found on the ninth to tenth day of pregnancy (Engle, 1927, 1931). Later, Cole (1937) obtained similar results with pregnant mare serum. Superovulation has been induced in several species by injecting gonadotropic preparations, but for the purpose of this discussion we are interested primarily in quantitative data that show the conditions under which the response can occur and the nature of the reaction. Unfortunately such papers are not many and data are difficult to compare due to differences in method of preparation and standardization of gonadotropic extracts. This is particularly true of pituitary gonadotropins for which there is no international unit.

Evans and Simpson (1940) used a highly purified FSH preparation in experiments on superfecundity in rats, as indicated by supernumerary implantation. This material was used on female rats 26 to 34 days of age that were placed with males 48 hours after the beginning of treatment. A total dose of 10 to 15 RU gave the highest number of implantation sites, while higher doses were progressively less successful and low doses inadequate. Almost all the animals that breed did so between 48 and 60 hours after the onset of treatment. These results are in agreement with those of Engle (pituitary implants) and Cole (PMS) and show that in both mice and rats approximately 30 implantations can be produced and the number of live young born (17) increased about three times the normal number.

Rowlands (1944) reports that 50 to 60 gram rats given single injections of 2 to 60 IU of pregnant mare serum did not give consistent results either in number of ova ovulated or number of animals responding. However, in a second series of experiments it was found that ovulation could be produced regularly in such animals if the serum treatment was followed by an injection of pregnancy urine hormone. The optimal interval between injection of serum gonadotropin and pregnancy urine hormone was 48 to 72 hours, with the best results at 56 hours. It is also interesting that the number of ova ovulated was correlated with the dosage of serum gonadotropin and not with the amount of pregnancy urine

hormone used to induce ovulation. Superovulation could be produced following the injection of 30 IU of serum gonadotropin but almost no ovulation could be obtained after a dose of 60 IU.

Williams (1945) has shown that a single injection of serum gonadotropin will produce ovulation in hypophysectomized rats until the fourth day following the operation, but not later. That the failure after the fourth day was due at least in part to ovarian atrophy was demonstrated by injecting serum gonadotropin followed by an injection of sheep or pig gonadotropin, pregnancy urine hormone, or a second treatment of pregnant mare serum (Rowlands and Williams, 1943). The optimum interval between treatments was found to be four days and fewer animals responded when the second treatment was postponed until the fiftieth day by which time follicular atrophy may have started.

If ovarian atrophy following hypophysectomy was prevented by implanting tablets of stilbestrol ovulation could be induced by a single injection of serum gonadotropin as in a normal animal (Williams, 1945).

Rabbits also have been used in investigations on superovulation. Pincus (1940) has reported extensive experiments on prepubertal and pubertal rabbits five or six months old and mostly of a chinchilla strain. His usual procedure was to stimulate follicular development by giving six subcutaneous injections of a gonadotropin over three days followed within eight to twelve hours by an intravenous injection to induce ovulation. The preparations used to promote the growth of follicles were pregnant mare serum, FSH fractions of sheep pituitaries, and horse pituitary extract. Those employed for intravenous injections were the serum gonadotropin, horse pituitary extract and a combination of FSH and LH extracts of sheep pituitaries. He found that all of these gonadotropic preparations would produce ovulation when properly administered but the pituitary extracts usually caused superovulation while mare serum would only occasionally do so. As many as 80 ova were obtained from a single rabbit. He also found that in these young animals copulation could not be substituted for the intravenous injection for inducing ovulation.

Parkes (1943) produced superovulation and superfecundation in adult rabbits by "priming" them with an extract of horse pituitaries and producing ovulation by an intravenous injection of pituitary extract or pregnancy urine hormone, or by mating. In experiments to determine the most favorable dosage and length of treatment for the growth of follicles in rabbits, he found 1 or 2 mgm. of the pituitary extract daily better than 5 mgm. and a five-day treatment better than three. The larger dosage produced cystic and hemorrhagic follicles. He also found that sexually immature Lop rabbits weighing about 2.0 kgm. would mate and become pregnant following a priming treatment. The ovaries of those rabbits weighing approximately 1.0 kgm. were insensitive to gonadotropins while animals between 1.0 and 2.0 kgm. showed two different conditions; the follicles responded to priming but did not ovulate after mating or ovulation took place but corpora lutea did not develop.

These investigations in which gonadotropins have been used to promote follicular development and ovulation, have a number of points in common. In

growing follicles with gonadotropins, there is both an optimal dosage and an optimal length of treatment. Comparisons of dosages for the different preparations are difficult to make, for reasons already mentioned, but the time required for the response is approximately the same for all. From the beginning of treatment, mating in rats occurs between 48 and 60 hours (Evans and Simpson, 1940); optimal time for inducing ovulation with pregnancy urine hormone in normal rats, 48 to 72 hours (best at 56 hrs.) and in hypophysectomized rats, four days (Rowlands, 1944; Rowlands and Williams, 1943). In rabbits, Pincus (1940) found three days as the optimal length of treatment for the immature animal while Parkes (1943) found five days better than three for the adult, both as to the number of ova ovulated and the number of large follicles in the ovaries. This is probably significant for the adult animal as Hill and White (1933) in a study of regression of follicles in the estrous rabbit noticed that a decrease in size and disappearance of follicles started on the fifth day.

These observations indicate that a follicle grows at a definite rate and the time required to reach a condition susceptible to ovulatory stimulation cannot be shortened greatly by increasing the growth stimulus. As the dosage of a gonadotropin is increased more and more follicles become involved in the growth process until a point is reached at which defects appear in the form of cystic follicles and premature luteinization. When pituitary extracts are used, such abnormalities may be due to the presence of LH either in the preparation or from the animal's own pituitary. Such defects are not found even when large doses of pure FSH are given to hypophysectomized rats (Greep, van Dyke and Chow, 1942). The luteinizing property of pregnant mare serum and pregnancy urine hormone also gives rise to a similar situation when they are employed as growth-promoting agents in follicular development.

Another point of interest is that the number of follicles that can be brought to ovulation by the priming process greatly exceeds the number of large follicles from which selection is made in the normal cycle prior to preovulatory enlargement. It seems obvious that a large number is included from those that had not arrived at the point of differentiation before the treatment was started. This does not imply necessarily that follicular growth rate was increased as it seems more probable that differentiation of the theca and granulosa was brought about before the follicle reached full size. This supposition has some support in the fact that ovulation can be induced before a follicle attains its full diameter (Dempsey, 1937).

Ovulation in sheep and cattle. The application of gonadotropins to problems in animal husbandry has contributed much information concerning ovulation in farm animals. Most of this work has been done in Russia, Great Britain, and the United States and the reader is referred to recent papers by Hammond (1945); Hammond and Bhattacharya (1943); and Cole, Hart and Miller (1945) for references to the literature. From the standpoint of this review the most useful data were obtained from studies of ovulation in sheep and cattle. Those for cattle are the most important due to the advantage that conditions in the ovaries can be determined by manual palpation per rectum (Hammond, 1927).

Cole and Miller (1933) showed that ovulation could be induced in anestrous sheep by injecting serum gonadotropin but the results were highly variable and not usually accompanied by estrus. Hammond (1945) mentions that the absence of estrus in such experiments is like the first ovulation of the natural breeding season and may be followed, by the length of an estrous cycle, by spontaneous ovulation and estrus. Also, a second treatment following the first by about sixteen days would cause ovulation and estrus. Probably the most suggestive observation by Hammond was that ovulation could be produced by injecting stilbestrol or stilbestrol di-n-butyrate. Seven out of eleven animals given 1 to 5 mgm. stilbestrol did not ovulate but ovulation occurred in five of six treated with 0.25 or 0.5 mgm. of stilbestrol di-n-butyrate. It also was found that stilbestrol tends to inhibit ovulation if given simultaneously or one or two days before serum gonadotropin, but when the estrogen was injected twelve to twenty-four hours after the serum, there was no inhibition or it was of a lesser degree. Another observation was, that the ovulating action of both stilbestrol and serum gonadotropin generally was inhibited by the presence of a functional corpus luteum.

Several investigators have produced ovulation and multiple pregnancies in cattle but the most recent and, in many respects, the most detailed study is that by Hammond and Bhattacharya (1943). Most of their experiments were performed on cows and heifers that had been selected for slaughter while a few were bred and left to calve. They succeeded in producing superovulation and multiple pregnancies but the object of their investigation was a practical one, the production of twin ovulations. For this purpose they concluded, from preliminary experiments, that 1500 IU of serum gonadotropin or about 100 mgm. of horse pituitary extract was the most favorable dosage to use for stimulating follicular growth. However, before discussing this point the influence of a functional corpus luteum on ovulation should be mentioned.

It is now well known that if a functional corpus luteum is expressed from the ovary of a cow, estrus and ovulation follow within about four days. The corpus luteum apparently does not prevent follicular growth in response to a gonadotropin but it does inhibit ovulation. This seems particularly true for reasonably small doses (1500 IU) but heavy dosage (5000 IU) may produce ovulation in the presence of the corpus luteum. Hammond and Bhattacharya took advantage of these facts and most of their experiments deal with the relation between dosage of gonadotropin and the time of removal of the corpus luteum. Various amounts of serum gonadotropin and horse pituitary extract were given, in most instances, from five days before to two days after the corpus luteum was removed. However, for present purposes it seems sufficient to mention the results obtained when 1500 IU of serum gonadotropin was used. Although these results were quite variable it seems that the most effective time for administering this dosage for producing approximately two ovulations was two days before to the day after corpus luteum removal.

These observations on sheep and cattle show that progesterone and estrogen have an important influence on ovulation. The presence of a functional corpus

luteum (presumably progesterone) inhibits ovulation. In sheep, estrogen seems to have two effects one of which depends upon the time it is injected in relation to the administration of a gonadotropin. Estrogen alone may induce ovulation if a large follicle is present and if injected before a gonadotropin may inhibit the ovulation that ordinarily occurs.

Influence of estrogen and progesterone on ovulation. Observations similar to those reported for sheep and cattle have been made on laboratory animals. From these studies it seems that ovulation produced by the injection of estrogen is not due to a direct action on the ovary but probably results from an effect on the pituitary causing an increased output of LH. It has been shown in rats that the secretion of LH by the pituitary can be stimulated by injecting estrogen (see Fevold, 1939, for references). A three-day treatment with FSH causes only follicular development in the ovaries of immature rats, but if preceded by three daily injections of estrogen both luteinization and augmentation of ovarian weight takes place. However, if hypophysectomized animals are used the estrogen treatment does not modify the response to FSH, thus showing that the LH came from the pituitary as a result of the action of estrogen.

The inhibitory action of estrogen on the response to serum gonadotropin in sheep also has its counterpart in experiments on rats (Fevold and Fisk, 1939). While a three-day treatment with estrogen will augment the response of the ovary to FSH, an eight-day treatment prior to the injection of the gonadotropin will greatly decrease the response. Fevold and Fisk found that rats on estrogen for eight days had a higher concentration of LH in the blood than those treated for three days. They also showed that the same results could be obtained by injecting an LH preparation, and demonstrated, in both normal and hypophysectomized animals, that the extent of inhibition was correlated with dosage of LH.

Thus it seems that while a small amount of LH augments the action of FSH the presence of a large amount, or effects produced by prior action, inhibits the response of the follicle to a follicle-stimulating gonadotropin. Yet, there is considerable uncertainty as to whether such inhibition is the result of a direct action of LH on the follicle or is brought about by some other means (Fraenkel-Conrat et al. 1940ab.; van Dyke, 1939). However, it is significant that pure LH will inhibit the gonadotropic action of pregnant mare serum on the ovaries of immature rats (Greep, van Dyke and Chow, 1942).

The inhibitory influence on ovulation exercised by the mammalian corpus luteum apparently is accomplished by the secretion of progesterone. It has been mentioned that a functional corpus luteum does not inhibit follicular growth but prevents preovulatory enlargement and rupture of the follicle. This is in line with experimental results indicating that progesterone inhibits secretion of LH by the pituitary. Makepeace, Weinstein and Friedman (1937) found that estrous rabbits under the influence of progesterone did not ovulate when mated. Astwood and Fevold (1939) likewise have shown that pseudopregnancy in rats, which usually follows electrical stimulation of the uterine cervix, did not occur if progesterone was injected at the time or prior to the application of

the stimulus. Progesterone also will prevent the action of estrogen on the pituitary resulting in the release of LH (Astwood and Fevold, 1939). It is a common observation that immature rats on a prolonged treatment with FSH show first follicular development followed by luteinization. Such luteinization is thought to be due to LH secreted by the pituitary in response to estrogen from the stimulated follicles. If, in such experiments, progesterone is also given the follicles are not luteinized. Therefore, it seems that when a gonadotropin which itself has no or only weak luteinizing capacity is used to grow follicles in the presence of a corpus luteum, as in cattle, ovulation does not take place because the luteinizing action of the injected hormone is insufficient or lacking and progesterone secreted by the corpus luteum inhibits release of LH from the pituitary. In this connection it is interesting to note that Hammond and Bhattacharya (1943) obtained no ovulation with 1500 IU of pregnant mare serum in ten cows having corpora lutea but of four others given 5000 IU one had a single ovulation.

There are, however, other facts that tend to suggest that progesterone may influence ovulation in ways other than those just mentioned. One of these is brought out by Hammond (1939) in a discussion on the treatment of cystic follicles in cattle. He found that after a cyst is ruptured another follicle develops but does not ovulate, but if this one is ruptured it luteinizes and is followed by a normal cycle and estrus three weeks later. Day (quoted by Hammond and Bhattacharya, 1943) found that a normal cycle may also be brought about by rupturing a cyst and administering pregnancy urine extract. These observations indicate that a period of luteal activity in some way restores the hormonal balance requisite for ovulation.

A very interesting series of studies of a similar nature was started by Everett in 1939 on a strain of rats that has a tendency to remain in constant estrus. This condition, once it appears, persists indefinitely except for short diestrous periods at infrequent intervals. The ovarian follicles are slightly cystic, being larger than normal, and the nuclei of the interstitial tissue present varying degrees of achromatolysis and "wheel" figures resembling the condition seen in hypophysectomized rats. These defects suggest a deficiency in LH while in the case of the follicles the mild overgrowth probably indicates very weak LH action.

Everett has found it possible to correct this situation and restore normal ovulation and luteinization by administering progesterone. In one series of experiments (Everett, 1940) he gave persistent estrous rats daily injections of progesterone in doses below that required to inhibit the estrous cycle of a normal animal. Several rats were kept on this treatment for 40 to 50 days and maintained fairly regular cycles, but all returned to constant estrus when the injections were stopped. Another method (Everett, 1943) was to give a single "interrupting" dose of 0.5 to 1 mgm. of progesterone which usually caused atresia of the large follicles and the appearance of diestrus lasting for two or three days, during which a new set of follicles grew. If the treatment was concluded at this point 90 to 95 per cent of the animals would return to persistent estrus following the brief diestrous interval. However, ovulation and luteinization could be induced

by a second injection if given at proper intervals after the first. The effectiveness of the second injection was slight in early diestrus, rose to a maximum on the first day of the succeeding estrus, and was negligible after estrus had persisted longer than eight days. Therefore, on the basis of these facts and the results of other similar studies (Everett, 1944a, b; 1945) Everett concluded that in rats small amounts of progesterone cause the release of LH from the pituitary.

Considerable emphasis has been placed on the relation of LH to ovulation but this does not mean necessarily that it is the hormone chiefly responsible for rupture of the follicle.

In all instances, so far mentioned, in which ovulation has been evoked experimentally the possibility has existed that both follicle-stimulating and luteinizing actions have been present. This has been brought about either by endogenous gonadotropins or by the physiological nature of the gonadotropin injected. This has been true even in experiments in which, after follicular development, ovulation has been induced in hypophysectomized animals by the injection of pregnancy urine hormone. This chorionic gonadotropin is not entirely a luteinizing substance as is so often stated in the literature (see Engle, 1939, for references). Therefore, it seems possible that pre-ovulatory enlargement and ovulation in the normal estrous cycle may be accomplished by a balanced action of the two pituitary hormones, FSH and LH, not one, often referred to as "the ovulating hormone."

Analysis of the actions of FSH and LH in ovulation. Evidence in support of the theory that both FSH and LH are required for ovulation can be drawn from investigations in which partially fractionated pituitary preparations have been used under conditions that minimized or eliminated participation of the animal's own pituitary. Such studies include ovulation in anestrous and immature cats (Foster and Hisaw, 1935), juvenile rabbits (Foster and Fevold, 1938) and hypophysectomized adult rabbits (Foster, Foster and Hisaw, 1937). Follicular development in all instances was produced by daily injections of FSH for five days and ovulation was induced by a single intravenous injection of FSH, LH, or a mixture of the two hormones. In both cats and juvenile rabbits it was found that ovulation could be evoked by an adequate dose of either FSH or LH but a mixture of the two was more effective, as judged by the number of ruptured follicles. It was also found that when the amount of FSH in the ovulating mixture was kept constant and various amounts of LH were added, ovulation decreased as the concentration of LH increased until a point was reached at which no ovulation took place. When similar preparations were tested on hypophysectomized rabbits it was found that neither FSH nor LH alone could produce ovulation but a combination of the two was effective.

Greep, van Dyke and Chow (1942) tested the ovulating action of pure FSH and LH on normal estrous and post partum rabbits. Both hormones produced ovulation and LH was found slightly more effective than FSH. They also tested the effects of combined injections of the two hormones and demonstrated a potentiation but of a smaller degree than that obtained by Foster, Foster and Hisaw (1937). It is unfortunate that these pure preparations were not tested

also on hypophysectomized rabbits. However, these experiments, as a whole, tend to support the theory that ovulation results from the combined actions of FSH and LH.

Probably the most promising material for an analysis of the process of ovulation is the amphibian ovary. Since the first experiments on the induction of ovulation by Wolf (1929) on *Rana pipiens* and Houssay, Giusti, and Lascano-Gonzalez (1929) on *Bufo marinus* similar observations have been made by many investigators on various species of both Anura and Urodela. Only a few papers have been selected that have direct bearing on the present discussion and the reader is referred to a review by Creaser and Gorbman (1939) for a general consideration of the literature.

It has been found recently that ovulation can be produced in *Rana pipiens* by administering fractionated gonadotropins from mammalian pituitary glands (Wright and Hisaw, in press). This has made it possible to repeat on the frog several experiments that have been performed thus far only on mammals. One of these is the action of FSH and LH in the hypophysectomized frog. It was found that an FSH preparation would produce ovulation in the normal frog but could not do so in the hypophysectomized animal. However, both the normal and hypophysectomized animal would respond to combinations of FSH and LH.

Another effect produced by FSH, first observed in the frog, is a sensitization of the ovary to an ovulating stimulus. Rugh (1939) found that hypophysectomy greatly increased the sensitivity of the ovary to implanted pituitary tissue. Wright (1946) confirmed this observation and suggested that the effect was probably due to a release of gonadotropins from the pituitary during the operation. This idea is in agreement with the observations of Lane and Greep (1935) and Williams (1945) who found that pituitary ablation produced stimulation of the ovaries in rats. In the frog, one of the hormones released is apparently FSH. The injection of a small amount of an FSH preparation, that is insufficient for ovulation, will increase the responsiveness of the ovary in a normal frog. Also, hypophysectomized frogs given eight daily injections of an FSH preparation, adequate to ovulate a normal animal, did not ovulate, but their ovaries remained more responsive to an ovulating mixture of FSH and LH than those of a normal animal. The ovaries of similar animals, not treated for the same period, become refractory thus indicating that an FSH preparation also prevents involutionary changes in the follicles that ordinarily result from loss of the pituitary.

An advantage offered by the frog's ovary is that it will ovulate in vitro and the process can be watched. Heilbrunn, Daugherty, and Wilbur (1939) produced ovulation in *Rana pipiens* by suspending one ovary in Ringer's solution containing a triturated pituitary gland of a frog. Ryan and Grant (1940) obtained ovulation by putting small pieces of ovary in vials containing a suspension of frog pituitary tissue in Ringer's solution. Wright (1945) has made a rather extensive study of the conditions influencing ovulation of the frog's ovary in vitro. His experiments were conducted on bits of ovary including 15 to 30 ova which were suspended in vials containing 10 cc. of Holtfreter's fluid. This made it possible to have a large number of duplicates and controls from the

same ovary. Ovulation was induced by adding various amounts of a fine suspension of frog pituitary tissue. Probably the most important observation, as it relates to the present topic, is the influence of concentration of pituitary suspension on the percentage of ovulation. This was tested by adding to a series of vials containing tissue from the same ovary, different amounts of the suspension ranging from that equivalent to two pituitaries to $\frac{1}{512}$ of a gland. The results for such experiments show a gradual increase in the percentage of eggs ovulated until a concentration of $\frac{1}{8}$ to $\frac{1}{16}$ of a pituitary is reached after which there is a progressive decline. In fact, in some tests no ovulation took place at concentrations equivalent to one or two pituitary glands. The reason for the inefficiency of the higher concentrations is not known but the observations on juvenile rabbits by Foster and Fevold (1938) suggest that it might be correlated with the increase in amount of LH as more pituitary tissue was added to the solution.

A partial answer to this question probably can be found in the results for in vitro experiments in which mammalian gonadotropins were used. Wright and Hisaw (in press) have found it possible to induce ovulation in bits of frog ovaries in vitro with extracts of sheep pituitary glands under the same conditions as those just described. They found that an ovarian fragment in Holtfreter's fluid could not be made to ovulate by adding an FSH preparation. However, the ovarian tissue became much more responsive to mixtures of FSH and LH than bits of the same ovary that had not been treated with FSH. For example, ovulation of ovarian fragments of a normal ovary was not obtained with mixtures more dilute than $\frac{1}{4}$ unit of LH plus $\frac{1}{2}$ unit of FSH but similar bits of ovary treated with FSH responded to as little as $\frac{1}{16}$ unit of LH plus $\frac{1}{32}$ unit of FSH. The optimal balance between the two hormones for ovulating either normal or sensitized bits of ovary was not determined but it was found that $\frac{1}{2}$ unit of LH plus $\frac{1}{6}$ unit of FSH was ineffective while $\frac{1}{2}$ unit of LH plus 1 unit of FSH produced ovulation. Also, the effect of LH alone was not determined due to the lack of a preparation of sufficient purity.

These investigations on the frog are of general interest as they indicate that probably the physiological processes concerned with ovulation are basically the same in all vertebrates. It must be significant that FSH preparations do not cause ovulation in hypophysectomized rabbits, hypophysectomized frogs, nor in bits of frog ovaries in vitro, while proper combinations of FSH and LH will do so in all instances. Yet it is a fact that the information we now have is based for the most part on results in which pituitary preparations of questionable purity have been used. Such data are no doubt reliable in most respects but should be accepted with reservation until the effect of chemically pure pituitary gonadotropins have been investigated.

Conclusions on ovulation. The results for experiments on ovulation tend to support the theory that rupture of the follicle is brought about by the synergistic action of FSH and LH. Yet the quantitative relationship between these two hormones, that is most favorable for ovulation, has not been determined for a single species. It also seems likely that there is considerable variation and that

each species, though similar to all others, has its individual pecularities. This certainly seems true when one considers the responsiveness of some and the difficulties encountered with others. With regard to this one need only mention the discouraging experiences of those who have attempted to produce ovulation in primates (Hartman, 1938; Davis and Hellbaum, 1944; Abarbanel and Leathem, 1945).

It is also a problem as to how the proper balance between the pituitary hormones, at the time of ovulation, is achieved. It seems to be the general opinion that this is accomplished by an increase in secretion or release of hormones from the pituitary as a result of nervous stimulation at mating or in response to the ovarian hormones, estrogen and progesterone. A point made in the discussion is that estrogen is found in all vertebrates while progesterone probably is present only in mammals and therefore cannot be of general importance in ovulation. Although Everett has presented strong evidence in support of the possibility that small doses of progesterone stimulate the secretion of LH by the pituitary in rats, it is also true that larger doses and functional corpora lutea inhibit ovulation. The writer is inclined toward the opinion that these actions of progesterone are mammalian adaptations which assist in regulating the estrous cycle while among the vertebrates as a whole the action of estrogen on the pituitary is probably more important as far as ovulation is concerned. Also, one might expect that the pituitary would not respond to estrogen in those species that require the nervous stimulation of mating for ovulation. Bachman (1935) found this so for the rabbit, but whether it is generally true is not known. On the other hand, the pituitaries of those species that ovulate spontaneously should react to estrogen. Evidence for this has been presented for the rat and sheep but such information for other mammals is almost entirely lacking.

If it is agreed that the output of gonadotropins by the pituitary is increased during ovulation there is yet the question as to whether it concerns FSH, LH or both. An increased secretion of LH is mentioned most frequently in the literature. These statements concerning LH are probably based on the close association of luteinization with ovulation but at present there are no quantitative data that are suitable for settling this point. The results of research on the rabbit are the nearest approach. Hill (1934) found that mating in rabbits was followed by a rapid decrease in the ability of the pituitary to induce ovulation when tested in estrous does and Westman and Jacobsohn (1937) found a corresponding increase in the ovulating capacity of the blood. These observations demonstrate a greater output of gonadotropin but they are not quantitative in the sense that they indicate the amount of the hormone or hormones involved. The writer is of the opinion that the evidence is in favor of LH but these results could be explained by assuming an increase in LH, FSH or both as either or a combination of the two will cause ovulation in a normal estrous rabbit. There is much to commend in the idea expressed by Dempsey (1937) that it is not necessary to postulate wide fluctuations in FSH during the estrous cycle as all the phenomena could be explained on the assumption that FSH remains quite constant and ovulation results from an increase in LH. This and many similar

problems could be solved if accurate quantitative techniques were available for determining the FSH and LH content of unfractionated preparations of small samples of blood and single pituitary glands of laboratory animals.

It also is clear that before a complete explanation of ovulation can be given, the relative importance of FSH and LH must be determined. At present we have no reliable quantitative data on this point. It is true that important and suggestive observations of a quantitative nature have been made, but in all instances reasonable doubt can be raised either with regard to the purity of the pituitary preparations used, or the possibility of endogenous gonadotropins being present. In experiments on mammals these objections can be met by using pure FSH and LH and hypophysectomized animals. It also is possible that very valuable experimental material for investigations on ovulation can be found among the cold-blooded vertebrates. Amphibia hold the most promise and offer the advantage of studying ovulation in vitro. We now have the basic information, techniques, and materials for solving this problem and one might expect significant progress within the near future.

SUMMARY

Follicular development. The development of a Graafian follicle can be divided into four stages. (1) The first includes oogenesis, organization of the granulosa and theca interna, and growth up to the formation of the follicular antrum. In mammals the ovum attains approximately its full size. A characteristic of this period is that the growth processes probably are regulated by a self-contained system of organizers. (2) The second stage begins when the follicle gains competence to respond to the action of pituitary gonadotropins. It is marked by rapid growth, mitotic activity in the theca interna and granulosa, and multiplication of follicular blood vessels. (3) The third stage is introduced by a rapid falling off of mitotic rate in both theca interna and granulosa. The blood vessels continue to increase and reach full development in the fourth stage. This seems to be a period of differentiation, indicated by hypertrophy of the cells in the theca interna and granulosa. Increase in size of the follicle is due mostly to the accumulation of fluid in the antrum. (4) The fourth stage includes pre-ovulatory enlargement of the follicle and ovulation. The chief features are the rapid secretion of follicular liquor, marked hyperemia of the follicle, and usually the ovum undergoes the first polar division.

The physiological processes of the last three stages proceed under the co-ordinated actions of the pituitary and the ovarian hormones.

Ovulation. Much regarding the physiology of ovulation remains obscure but the results of recent research warrant the following tentative conclusions: (1) That pre-ovulatory enlargement and rupture of the follicle is produced by the joint action of FSH and LH. (2) That neither FSH nor LH acting alone can cause ovulation. (3) Ovulation is initiated by an increase in the secretion of gonadotropins by the pituitary and probably this mostly concerns LH. (4) This increased output of gonadotropins is caused, in species that ovulate spontaneously, by the action of estrogen and possibly progesterone on the pituitary, and

probably by nervous stimulation of the pituitary in those species that require mating for ovulation. (5) Another possibility worth considering is that in certain species the sensitizing effect of FSH on the follicle (frog) may assist the ovulatory reaction or even make it possible for the follicle to respond to the existing concentration of gonadotropins thus making an increase unnecessary for ovulation.

REFERENCES

ABARBANEL, A. R. AND J. H. LEATHEM. Am. J. Obst. and Gynec. 50: 262, 1945.

ALLEN, E., F. L. HISAW AND W. U. GARDNER. Sex and internal secretions. Chapt. VIII 2nd ed. Williams & Wilkins, 1939.

ASCHHEIM, S. Ztschr. f. Geburtsh. u. Gynäk. 90: 387, 1926.

ASTWOOD, E. B. AND H. L. FEVOLD. Am. J. Physiol. 127: 192, 1939.

ASTWOOD, E. B. AND R. O. GREEP. Science 89: 81, 1939.

ATWELL, W. J. Anat. Rec. 55: Suppl. p. 45, 1933.

BACHMAN, C. Proc. Soc. Exper. Biol. and Med. 33: 551, 1935.

BOLING, J. L., R. J. BLANDAU, A. L. SODERWALL AND W. C. YOUNG. Anat. Rec. 79: 313, 1941.

BRAMBELL, F. W. R., U. FIELDING AND A. S. PARKES. Proc. Roy. Soc. London, S.B. 102: 385, 1928.

BULLOUGH, W. S. Nature 149: 271, 1942.

J. Endocrinology 3: 211, 1942.

J. Endocrinology 3: 150, 1942.

BURNS, R. K. Anat. Rec. 51: 333, 1932.

CHOW, B. F. Annals New York Acad. Sci. 43: 309, 1943.

CHOW, B. F., H. B. VAN DYKE AND R. O. GREEP. Endocrinology 30: 650, 1942.

CLAUBERG, C. Ztschr. f. Gynäk. 16: 964, 1932.

COLE, H. H. Am. J. Physiol. 119: 704, 1937.

COLE, H. H., G. H. HART AND R. F. MILLER. Endocrinology 36: 370, 1945.

COLE, H. H. AND R. F. MILLER. Am. J. Physiol. 104: 165, 1933.

CORNER, G. W. Physiol. Rev. 18: 154, 1938.

CREASER, C. W. AND A. GORBMAN. Quart. Rev. Biol. 14: 311, 1939.

DAVIS, M. E. AND A. A. HELLBAUM. J. Clin. Endocrinology 4: 400, 1944.

DAWSON, A. B. Anat. Rec. 79: 155, 1941.

DEMPSEY, E. W. Am. J. Physiol. 120: 126, 1937.

DEMPSEY, E. W. AND D. L. BASSETT. Endocrinology 33: 384, 1943.

DEMPSEY, E. W., R. HERTZ AND W. C. YOUNG. Am. J. Physiol. 116: 201, 1936.

DRIPS, D. G. AND F. A. FORD. Surg., Gynec. and Obst. 55: 596, 1932.

ENGLE, E. T. Proc. Soc. Exper. Biol. and Med. 25: 85, 1927.

Endocrinology 15: 405, 1931.

Anat. Rec. 48: 341, 1931.

Sex and internal secretions. Chapt. XVIII, 2nd ed. Williams & Wilkins, 1939.

EVANS, H. M. AND M. E. SIMPSON. Endocrinology 27: 305, 1940.

EVERETT, J. W. Endocrinology 25: 123, 1939.

Endocrinology 27: 681, 1940.

Endocrinology 32: 285, 1943.

Endocrinology 34: 136, 1944a.

Endocrinology 35: 507, 1944b.

Am. J. Anat. 77: 293, 1945.

FEVOLD, H. L. Cold Spring Harbor Symposia on Quant. Biol. 5: 93, 1937.

Sex and internal secretions. Chapt. XVII, 2nd ed. Williams & Wilkins, 1939.

Annals New York Acad. Sci. 43: 321, 1943.

FEVOLD, H. L. AND V. M. FISKE. Endocrinology 24: 823, 1939.

Foster, M. A. and H. L. Fevold. Am. J. Physiol. **121**: 625, 1938.
Foster, M. A., R. C. Foster and F. L. Hisaw. Endocrinology **21**: 249, 1937.
Foster, M. A. and F. L. Hisaw. Anat. Rec. **62**: 75, 1935.
Fraenkel-Conrat, H., C. H. Li and M. E. Simpson. Essays in biology. Univ. of Calif. Press, 1943.
Fraenkel-Conrat, H., C. H. Li, M. E. Simpson and H. M. Evans. Endocrinology **27**: 793, 1940.
Fraenkel-Conrat, H., M. E. Simpson, C. H. Li and H. M. Evans. An. Fac. de Med. de Montevideo **25**: nos. 5, 6, 7, 8, 1940.
Greep, R. O., H. B. van Dyke and B. F. Chow. Endocrinology **30**: 635, 1942.
Hammond, J. The physiology of reproduction in the cow. Cambridge, 1927.
Vet. Rec. **31**: 1137, 1939.
Hammond, J., Jr. J. Endocrinology **4**: 169, 1945.
Hammond, J., Jr. and P. Bhattacharya. J. Agric. Science **34**: Part 1, 1943.
Hargit, G. T. J. Morph. and Physiol. **49**: 277, 1930.
Hartman, C. G. Bull. Johns Hopkins Hosp. **63**: 351, 1938.
Sex and internal secretions. Chapt. IX, 2nd ed. Williams & Wilkins, 1939.
Heilbrunn, L. V., K. Daugherty and K. M. Wilbur. Physiol. Zool. **12**: 97, 1939.
Hertz, R. and F. L. Hisaw. Am. J. Physiol. **108**: 1, 1934.
Hill, R. T. J. Physiol. **83**: 129, 1934.
Hill, R. T. and A. S. Parkes. Proc. Roy. Soc. B. **116**: 221, 1934.
Hill, M. and W. E. White. J. Physiol. **80**: 174, 1933.
Hisaw, F. L. Yale J. Biol. and Med. **17**: 119, 1944.
Hohlweg. W. Klin. Wchnschr. **13**: 92, 1934.
Hohlweg, W. and A. Chamarro. Klin. Wchnschr. **16**: 196, 1937.
Hooker, C. W. Endocrinology **30**: 77, 1942.
Houssay, B. A., L. Giusti and J. M. Lascano-Gonzalez. Rev. de la Soc. Argentina **5**: 397, 1929.
Lane, C. E. Anat. Rec. **61**: 141, 1935.
Am. J. Physiol. **110**: 681, 1935.
Lane, C. E. and F. R. Davis. Anat. Rec. **73**: 429, 1939.
Lane, C. E. and R. O. Greep. Anat. Rec. **63**: 139, 1935.
Li, C. H., M. E. Simpson and H. M. Evans. J. Am. Chem. Soc. **64**: 367, 1942.
Loeb, L. Deutsch. Med. Wchnschr. **37**: 17, 1911.
Makepeace, A. W., G. L. Weinstein and M. H. Friedman. Am. J. Physiol. **119**: 812, 1937.
Mossman, H. W. Am. J. Anat. **61**: 289, 1937.
Anat. Rec. **70**: 643, 1938.
Parkes, A. S. Proc. Roy. Soc. London, S.B. **100**: 72, 1926.
Proc. Roy. Soc. London, S.B. **102**: 51, 1927.
J. Endocrinology **3**: 268, 1943.
Pencharz, R. I. Science **91**: 554, 1940.
Pincus, G. Anat. Rec. **77**: 1, 1940.
Price, D. Am. J. Anat. **75**: 207, 1944.
Physiol. Zool. **17**: 377, 1944.
Price, D. and E. Ortiz. Endocrinology **34**: 215, 1944.
Rowlands, I. W. J. Endocrinology **3**: 384, 1944.
Rowlands, I. W. and P. C. Williams. J. Endocrinology **3**: 310, 1943.
Rugh, R. Proc. Soc. Exper. Biol. Med. **40**: 132, 1939.
Ryan, F. J. and R. Grant. Physiol. Zool. **13**: 383, 1940.
Schmidt, I. G. Am. J. Anat. **71**: 245, 1942.
Schmidt, I. G. and F. G. Hoffman. Am. J. Anat. **68**: 263, 1941.
Shapiro, B. G. and H. A. Shapiro. J. Exper. Biol. **11**: 73, 1934.
Simpson, M. E., H. M. Evans, H. L. Fraenkel-Conrat and C. H. Li. Endocrinology **28**: 36, 1941.

SMITH, P. E. Proc. Soc. Fxper. Biol. and Med. 24: 131, 1926.
J. A. M. A. 88: 158, 1927.
Am. J. Anat. 45: 205, 1930.
Sex and internal secretions. Chapt. XVI, 2nd ed. Williams & Wilkins, 1939.
SMITH, P. E. AND E. T. ENGLE. Am. J. Anat. 40: 159, 1927.
SWEZY, O. Ovogenesis and its relation to the hypophysis. Science Press, 1933.
VAN DYKE, H. B. The physiology and pharmacology of the pituitary body. Univ. of Chicago Press, Vol. 2, 1939.
WESTMAN, A. AND D. JACOBSOHN. Acta Obst. et Gynaec. 17: 235, 1937.
WILLIAMS, P. C. Nature 145: 388, 1940.
Proc. Roy. Soc. B. 132: 189, 1944.
J. Endocrinology 4: 127, 1945.
J. Endocrinology 4: 131, 1945.
WITSCHI, E. J. Exper. Zool. 75: 313, 1937.
Sex and internal secretions. Chapt. IV, 2nd ed. Williams & Wilkins, 1939.
WOLF, O. M. Proc. Soc. Exper. Biol. and Med. 26: 692, 1929.
WRIGHT, P. A. J. Exper. Zool. 100: 565, 1945.
Physiol. Zool. (in press), 1946.
WRIGHT, P. A. AND F. L. HISAW. Endocrinology (in press), 1946.
ZONDEK, B. Ztsch. f. Geburtsh. u. Gynäk. 90: 372, 1926.

THE RENAL ORIGIN OF HYPERTENSION

HARRY GOLDBLATT

Institute of Pathology, Western Reserve University, Cleveland, Ohio

There is no reasonable doubt that some cases of human hypertension are of renal origin. The unsolved problem is whether many or most of those still referred to as "essential" are also of renal origin. The purpose of this review is to supplement the one made in 1940, in this Journal, by Blalock (24) and to summarize and evaluate the pertinent evidence that has now accumulated for or against the renal origin of human essential hypertension. Only relatively recent original references will be given here; for many of the established facts reference will be made to the bibliographies of previous reviews.

In the past, the renal origin of the elevated blood pressure in human essential hypertension has been excluded, by definition. Essential hypertension has usually been defined as persistent elevation of systolic and diastolic pressures, without accompanying renal disease, as evidenced by impairment of renal excretory function, when studied by the usual methods. Dalton and Nuzum (74) believe that critical statistical analysis of data on renal function in cases of human essential hypertension does show, in most cases, some reduction in the concentration of the urine and excretion of phenolsulphonephthalein. Yet it is a fact, known for many years, that in patients with long-standing hypertension, without obvious manifestations of renal excretory impairment during life, intrarenal arterial and arteriolar sclerosis, or even true nephrosclerosis, frequently may be found at autopsy (16, 109, 219, 254, 273). The unanswered questions are: (1) whether the hypertension or the vascular disease comes first, and (2) whether there is a causative relationship between the two conditions. Even to-day, as was the case 50 years ago (3, 214), some authors (153, 235, 306, 384) still regard the hypertension as primary, and of unknown origin, and the arterial and arteriolar sclerosis, both intrarenal and systemic, as independent of (3) or secondary to (214) the prolonged mechanical stretching of the walls of the blood vessels. Probably the only reason why some of these authors (153) concede that the hypertension of glomerulonephritis, pyelonephritis, and polycystic disease of the kidneys is induced by the renal disease is because it is known that these diseases of the kidney precede the hypertension. It is therefore not reasonable to contend that the hypertension is the cause of these inflammatory or developmental renal lesions. It is possible, however, to contend that the high blood pressure, or rather the increased bursting tension of the blood vessels, may be a mechanical cause of both intrarenal as well as systemic arteriosclerosis; therefore, this view has many modern adherents (153, 235, 274, 306, 384). These authors do not concede that the hypertension is a consequence of the renal vascular disease, although they agree that the renal disease may play a secondary part in maintaining the hypertension which is initiated either by a nervous (306) or a humoral (153) non-renal mechanism. Smith, Goldring and Chasis (384) regard the kidney as "the victim, rather than culprit" and

Schroeder (363) believes that essential hypertension may be the result of the combination of a number of factors.

If there is any one thing that is definitely established about the mechanism of the elevation of blood pressure, it is that widespread increase of peripheral (including splanchnic) vascular resistance is the determining factor (387) and that it is the tone of the systemic arterioles that is the determinant of peripheral vascular resistance. The old idea of v. Basch (3) that the arteriolosclerosis can be so generalized as to be a mechanical, organic cause of increased peripheral vascular resistance still persists, but no one has ever demonstrated such widespread disease of the blood vessels. One of the striking things about arteriosclerosis in man is how restricted the vascular disease is to certain organs and parts of the body, and especially how common it is in the kidneys (16, 109, 219, 254, 273). In this connection, the statement recently made (153) that "the kidneys are rarely involved to a significant degree in hypertensive disease; only about 8 per cent of all patients die of uremia", is misleading, for it attaches too much importance to renal excretion and ignores the important fact that the kidneys may be the seat of advanced vascular disease without significant disturbance of renal excretory function detectable by the most recent methods (108, 112). Finally, there is the view of the author (144) that the arterial and arteriolar sclerosis are primary, but of unknown origin, and that when the vascular disease affects the kidneys to a degree sufficient to produce adequate disturbance of intrarenal hemodynamics, a mechanism is initiated which brings about the increased peripheral vascular resistance that determines the hypertension. This view implies, of course, that any other pathologic process in the kidney which is capable of inducing the same disturbance of intrarenal hemodynamics would also be capable of initiating the mechanism that results in hypertension (76, 110, 202, 203, 206, 428).

It is obviously impossible to solve a problem of this kind by the study of patients, clinically, or by the examination of specimens of tissues obtained at autopsy. The arteriolar sclerosis of the kidneys, which has been reported by many authors (16, 109, 219, 254, 273), as an almost invariable finding in cases of human essential hypertension, at autopsy, is interpreted by some as proof for the view that the hypertension comes first and produces the renal vascular disease (44, 397). The possible causative relationship between the two conditions could conceivably be investigated by the examination of biopsy specimens of kidney obtained from the same individuals, at intervals, before and after the development of the hypertension. Since there is no way of telling with certainty which persons will develop hypertension, such a study of human beings is not feasible. An approach to this problem has been made by the examination of biopsy specimens of kidneys in cases of established human hypertension, after the condition had been in existence for various lengths of time. This study by Castleman and Smithwick (44, 397) has not led to a solution of the problem, for obvious reasons; but it is well to examine what it has contributed.

In a series of patients with hypertension that had existed for variable periods, biopsy specimens of the kidneys were examined for vascular changes. In these

small specimens in which, at best, only a few cross sections of independent arterioles could be observed, the changes were classified arbitrarily from 0 to ++++ on the basis of the amount of hyaline thickening of the wall and stenosis of the lumen. Despite the views and conclusions of the authors, it is an interesting and indeed an extraordinary fact that some vascular disease was found in 92 per cent of the specimens and that in about 50 per cent it was from moderate to severe. From the study of such a small specimen it is obviously impossible to tell whether the vascular disease is more or less pronounced in the remainder of the kidney. It is hazardous, therefore, to extrapolate from the evaluation of such a minute specimen of heterogeneous tissue to an estimate of the average condition of a single constituent, the arterioles, of the entire organ. An important fact, frequently overlooked, is that the larger intrarenal arteries are rarely included in such biopsy specimens, because the specimens are taken from the peripheral portion of the cortex. Even if the appearance of a few arterioles can give some estimate of the average change in similar vessels of the remainder of the kidney, yet it can give no idea of the state of the larger intrarenal vessels, or of the hemodynamic state of the entire kidney. It should be obvious that stenosis of one large intrarenal artery could account for great hemodynamic disturbance in a large mass of kidney, supplied by many arterioles that are not themselves diseased. Arteriosclerosis, with stenosis, of the larger *intrarenal* vessels is a common accompaniment of renal arteriolosclerosis. The possible contribution of stenosing sclerosis of the larger intrarenal arteries, and even of the main extrarenal artery, to the disturbance of intrarenal hemodynamics has been underestimated. Blackman (23) has shown how common it is to find at autopsy some degree of stenosis even of the main renal artery in hypertensive patients. At no time has it been asserted that stenosis of the main renal artery, unilateral or bilateral, is a common cause of, or frequent finding in, cases of human essential hypertension. This is why it is difficult to understand why Oppenheimer, Klemperer and Moschkowitz (287), and Lisa, Eckstein and Solomon (254), should make a special point of emphasizing that in many of their cases of hypertension in which pronounced stenosis of one or both main renal arteries was found, there was also significant intrarenal arteriolosclerosis. This is admitted freely by the author.

Experimental Renal Hypertension. The working hypothesis of the initial attempt (152) to produce experimental hypertension that might simulate human benign essential hypertension unaccompanied by disturbance of renal excretory function was based upon the assumption that, if the hypertension be the result of the *intrarenal* arterial and arteriolosclerosis, then the real cause of the elevated blood pressure might be the functional disturbance of renal hemodynamics produced by the stenosing vascular disease in the kidneys. It was considered that, if this were true, then the experimental production of such a disturbance of intrarenal hemodynamics by any means, even by constriction of the main renal arteries, by means of a clamp, might be followed by the development of hypertension without accompanying disturbance of renal excretory function. The clamp for the constriction of the main renal artery was not devised in order

to reproduce the effect of arteriosclerosis, with stenosis, of the main renal artery alone. It was fully recognized that this condition is not commonly found in patients with essential hypertension, although it has been shown to occur not infrequently (23). The clamp was developed as a device with which to reproduce experimentally, in animals, the probable hemodynamic disturbance in the kidney produced by *intrarenal* arterial and arteriolar sclerosis, so frequently found in persons with essential hypertension, or any other of a variety of pathologic processes, including stenosis of the main renal artery, which could produce a similar intrarenal circulatory disturbance. On the basis of an anatomical study of kidneys in human hypertensive persons Yuile (446) later found that intrinsic or extrinsic obstruction of the main renal artery simulating the renal clamp does occur, and concluded "that no evidence opposed to the view that experimentally produced hypertension has a human counterpart was encountered".

That hypertension which is unquestionably of renal origin, and which simulates human hypertension associated with renal vascular disease, has been produced experimentally, by constriction of the main renal artery, is now beyond question, for this has been fully confirmed by many investigators (92, 99, 102, 161, 163, 178, 195, 230, 293, 352), some of whom have employed modifications or different devices for the purpose of producing similar intrarenal hemodynamic disturbance.

In this review the account of the production of experimental renal hypertension will be made brief, because there have been many previous publications (146, 148, 149) about this phase of the subject. The original production (152) of experimental renal hypertension was accomplished by permanent constriction (not occlusion) of the main renal artery by means of a special silver clamp devised for the purpose. The effect of the clamp was to restrict the inflow of blood to the kidney, with resultant disturbance of intrarenal hemodynamics, the exact nature of which has not yet been definitely established.

There is no doubt that the immediate effect of constriction of the main renal artery is to restrict the flow of blood through the kidney (152, 251). Whether this effect persists indefinitely has not yet been determined with certainty (368). Restriction of the outflow of blood from the kidney by partial constriction of the main renal vein is also followed by transient hypertension (17, 118). It has been shown by Mann and collaborators (262) that it is necessary to reduce the caliber of an artery considerably before an appreciable effect on blood flow through the vessels is produced. All attempts to produce widespread intrarenal circulatory disturbance by emboli of various types had failed to produce hypertension (9, 42, 376). Reproduction of the widespread intrarenal arteriolar sclerosis would have been desirable, but it had not yet been accomplished at that time and it has not yet been produced in the dog. It is only recently that the production of so-called nephrosclerosis in chicks and rats has been reported by Selye and collaborators (371, 372, 373, 374, 375). There is no convincing proof in any of these studies that widespread intrarenal arteriolar sclerosis has been produced. The changes described are mainly parenchymatous degeneration and interstitial fibrosis, with an occasional arteriole showing fibrosis,

hyalinization and even necrosis. Some of these animals developed elevated blood pressure (372). The production of nephrosclerosis even in the chick and rat has not yet been confirmed; but others (38) have reported the hypertensive effect on the rat of repeated injections of desoxycorticosterone acetate.

Whether the results obtained by means of constriction of the main renal arteries can be applied completely to the problem of human essential hypertension still remains unanswered; but there is adequate indication that experimental renal hypertension and human essential hypertension are closely similar, if not identical, and that the results obtained in studies of pathogenesis, prevention, or treatment of the one may be applied directly to the other.

There are, however, those (153, 235, 254, 287) who refuse to assign great significance to these experiments, because, in human hypertension, the main renal artery is not commonly stenotic and therefore the vascular disease of the human kidney is not identical with experimental constriction of the main renal artery. What is forgotten, or ignored, by these authors is that constriction of the main renal artery was an expedient resorted to experimentally because it was the only available method whereby to produce a disturbance of the intrarenal circulation that might simulate the most probable effect of intrarenal stenosing arterial and arteriolar sclerosis. To think that the experimental type of hypertension is not exactly like human benign or malignant essential hypertension, because the main renal artery in human beings with hypertension is rarely stenotic (254), is to misunderstand the whole problem and the main purpose of the procedure which was used for the production of experimental renal hypertension. Such a concession as that made by Goldring and Chasis (153), namely, that a mechanism analogous to that of experimental renal hypertension can initiate hypertension in man but that this occurs but rarely, is gratuitous, and again ignores the probable similarity between the effect of widespread *intrarenal* vascular disease and experimentally produced stenosis of a main renal artery, which was the basis of the original experimental investigation (152). Recognition of this similarity is essential to a proper understanding and evaluation of the contribution made by the production of experimental hypertension by constriction of the main renal artery.

Temporary Hypertension Due to Constriction of One Main Renal Artery. Temporary hypertension which gradually returns to normal, in about four to six weeks, has been produced by constriction of one main renal artery. If the clamp is removed, and the circulation of the kidney restored to normal, or if the one kidney with the renal artery constricted is excised, the blood pressure returns to normal in 24 hours or less (25, 146, 223). In some animals, especially rats, goats, sheep, and even an occasional dog, unilateral constriction of the main renal artery is followed by the development of hypertension which persists for many months. Even in such animals the blood pressure returns to normal in 24 hours, or less, if the kidney with constricted renal artery is excised. It has been shown that, in dogs at least, the blood pressure may return to normal in as little as six hours.

In addition to affording evidence in favor of the renal origin of this type of

experimental hypertension, the significance of these observations on animals for the problem of hypertension in man has been the recognition of the existence of persistent hypertension associated with unilateral renal disease in man and of a possible cure of the hypertension, by excision of the diseased kidney, provided the other kidney is normal. The subject of the existence of hypertension associated with unilateral disease in man, and the return of the blood pressure to normal, in some patients, as the result of the removal of the diseased kidney, has already been reviewed (1) and more cases, some with return of blood pressure to normal after unilateral nephrectomy, have now been reported (10, 11, 12, 18, 27, 29, 30, 124, 141, 202, 255, 270, 316, 317, 341, 344, 346, 354, 376, 377, 379, 381, 384, 392, 394, 399, 427, 428, 429, 444, 445).

There is no convincing proof that intermittent total renal ischemia results in permanent elevation of blood pressure (148). There is every reason to believe that those who still insist that this does occur, in an exceptional case, were really dealing with permanent disturbance of renal circulation in an occasional animal (5, 257).

Persistent Hypertension Due to Constriction of Both Main Renal Arteries, and other Measures. Persistent hypertension has now been produced in dogs, monkeys, rabbits, sheep and goats (146), but not in the frog (408), by the constriction of both main renal arteries (146). As in the case of unilateral constriction of the main renal artery, a significant rise of blood pressure does not usually manifest itself in less than 24 hours, and often longer, although there is some evidence that the hypertensive effect may at least be initiated soon after the constriction (405). Hypertension due to constriction of one main renal artery may be made permanent by subsequent excision of the contralateral normal kidney (144, 225), by constriction of the main renal artery of the contralateral kidney (152), by occlusion of the ureter of the contralateral kidney (147, 271), or by the enclosure of one or both kidneys (preferably decapsulated) in a fishskin condom (146, 149). This type of membrane does not induce the development of a thick hull of connective tissue around the kidney, such as is induced by cellophane, collodion or silk (294, 296), but it does prevent the development of accessory circulation to the kidney. If the constriction of both main renal arteries is made moderate, the hypertension which results is not accompanied by detectable disturbance of renal excretory function—the benign phase (60, 152). If the constriction of both main renal arteries is made great, at the first operation, there is likely to be definite disturbance of renal excretory function, and fatal uremia may develop—the malignant phase (145, 147, 443). That complete renal ischemia for a few hours may result in anatomical and functional damage has been shown by Selkurt (370).

In those animals in which the blood pressure returns to normal after unilateral or bilateral constriction of the main renal arteries, the examination of the kidneys frequently shows that the naturally abundant, but poorly developed, accessory circulation to the kidney has become strikingly prominent, with large arterial vessels entering the cortex of the kidney from various surrounding sources. This observation indicates the possibility of treating experimental renal hy-

pertension by inducing the development of the accessory circulation to the kidney. Although the hypotensive effect of this procedure has been demonstrated experimentally, in dogs, by decapsulation and nephro-omentopexy (45) and by implantation of a part of the kidney into the spleen (spleno-renopexy) (425), yet this has not proved a fruitful lead for the treatment of human hypertension.

Pathologic Changes in the Organs of Animals with Persistent Hypertension. In the benign phase, in animals, even after six years of persistent hypertension, no significant pathologic changes develop in the aorta or in the large or small arteries. The only changes observed in the vascular system of these animals are slight to moderate hypertrophy of the heart (53) and thickening of the media of the large and small arteries (144). The results of these experiments, therefore, afford no proof for the view that hypertension *per se* is a sufficient condition for the production of generalized arterial or arteriolar sclerosis in the animals used. The adenomatous hyperplasia of the adrenal cortex reported in cases of essential human hypertension (345) has not been confirmed by others and has not been reported in animals with experimental renal hypertension.

In the malignant phase in animals, even when it terminates fatally in as little as 48 to 72 hours, the most profound changes occur in the blood vessels of the dog (147) or rabbit (322, 439) similar to those of the terminal phase of human malignant hypertension. In the aorta there is only interstitial edema, mainly of the media, but in the small arteries and arterioles, necrosis and fibrinoid degeneration occur, with or without acute inflammation in and around the wall of the vessels. The arteriolonecrosis and necrotizing arteriolitis, observed in many organs, are indistinguishable from similar changes observed in the terminal phase of malignant human hypertension.

The vascular changes of the malignant phase do not occur in the blood vessels of bilaterally nephrectomized animals that develop profound azotemia, without elevation of blood pressure, or in animals with great degree of hypertension of long standing, without renal excretory functional disturbance. They do not occur, nor does hypertension develop, in dogs in which the urine is deviated into the blood (139). In the malignant phase, in animals, the changes do not occur in the vessels of the kidney beyond the site of the clamp, where there is no hypertension, and where hypotension probably exists. For these reasons, it has been considered that a combination of impairment of renal function and increased bursting tension of the vascular wall are necessary conditions for the production of the pathologic changes in the vessels and that, even in human beings, these changes usually occur only terminally. The idea that these changes are induced by long-standing (months to years) intense vasospasm (409) is not upheld by these experiments. The inflammation around the blood vessels is probably only a reaction to the degeneration and necrosis of the walls of the arterioles, but, of course, may be caused by the same agent that produces the necrosis. These pathologic changes in the arterioles are not to be confused with arteriolosclerosis and are certainly not to be used as proof that hypertension itself can produce true arterio- or arteriolosclerosis. This erroneous interpretation has been made (90, 274).

We have not been able to confirm the observation (242, 243, 442) of vascular lesions of the malignant phase in animals that have been injected even with large quantities of renin, subcutaneously, intramuscularly or intravenously. This may be due to the difference in the method of preparation and purity of the renin employed.

Some investigators have produced lesions of arterioles which they consider similar to those of the malignant phase, by repeated intravenous injections of tyramine (93), but there is no good reason for believing that tyramine is responsible for the lesions in malignant hypertension.

The changes in the kidneys themselves depend upon the degree of constriction of the main renal artery and may vary from no detectable change, except in the mitochondria of the proximal convoluted tubules (101), to partial or even complete necrosis, or atrophy and fibrosis of the kidney (152). The conclusion of Wilson and Byrom (436, 437, 438) confirmed by others (122) that hypertension *per se* can produce the characteristic malignant arteriolar lesions in the rat, with hypertension due to constriction of one main renal artery, has not been substantiated in experimental renal hypertension in the dog, or any animal other than the rat (144, 147), and is probably based on the erroneous assumption that the contralateral kidney was normal. All the glomerular and interstitial inflammatory lesions described by these authors in the contralateral, supposedly normal kidney, have been observed by the author in rats with normal blood pressure. The author's explanation of the changes in the arterioles of the contralateral kidney, observed by Wilson and Byrom, and the others, is that they were unaware of the frequent existence of hydronephrosis or pyelonephritis in one or both kidneys of adult rats, and that, by constriction of one main renal artery, they were dealing, in some of these rats, with bilateral renal disease.

Pathogenesis of Experimental Renal Hypertension. This has been the subject of many publications, since the production of experimental renal hypertension was announced (152). Most of these accounts have dealt with the possible application of the observations on animals to the problem of the pathogenesis of human essential hypertension (24, 25, 140, 148, 208, 224, 235, 236, 271, 279, 280, 303, 306, 378, 384).

Progress in the acquisition of knowledge about the pathogenesis of human hypertension has been delayed because, in human beings, it is rarely if ever possible to study the patient before and after the development of the elevated blood pressure. The production of the counterpart of the benign and malignant phases of human hypertension has made possible the investigation of the probable pathogenesis of the high blood pressure, even in man. As soon as this was accomplished, the possible mechanisms involved in the elevation of the blood pressure became obvious: 1, afferent stimuli from the nerve endings in the kidneys to the vasomotor center or sympathetic ganglia, with resultant generalized vasoconstriction and consequent elevation of blood pressure; 2, afferent stimuli from the kidneys, with resultant output of an increased amount of some internal secretion which produces vasoconstriction either by central or peripheral action; 3, the accumulation or new formation of some substance, probably of renal origin, or a disturbance of equilibrium between substances already present

in the blood, resulting in a pressor effect like that of some of the known pressor substances.

Neurogenic Mechanism. In rapid succession the possible part played by various portions of the nervous system was investigated. It was soon shown that experimental renal hypertension is not caused by a nervous reflex from the kidney, affecting the vasomotor mechanism of the body. Renal denervation (53, 293); bilateral supradiaphragmatic excision of splanchnic nerves and lower dorsal sympathetic ganglia (148); subdiaphragmatic splanchnicectomy, with excision of the celiac and upper lumbar ganglia (25); bilateral section of the anterior nerve roots, from the sixth dorsal to the second lumbar inclusive (148); pithing (destruction of the entire spinal cord) (49, 143) and complete sympathectomy, including denervation of the heart (7, 117, 199) neither prevent, nor abolish the hypertension which results from the constriction of the main renal artery. Under pentothal anesthesia, experimental renal hypertension persists, whereas known neurogenic hypertension disappears (305). These results eliminated a nervous reflex originating in the kidney as the cause of the hypertension and indicated that a humoral mechanism was probably the cause. Ogden (283) however, who concedes that the renal humoral pressor mechanism initiates experimental renal hypertension, believes that he has demonstrated, by experiments on rats, that it is superseded later by a neurogenic mechanism mediated through the sympathetic nervous system. More work on this phase of the subject is required.

Although there are those who still cling to the idea that essential hypertension is endocrinogenic, and due to some disturbance of the hypophysis, yet the experiments on animals fail to confirm this. Hypophysectomy (8, 103, 148, 287, 315) has no definite effect. There is no definite proof (114) that the activity of the posterior lobe of the pituitary is altered in experimental renal hypertension. There is no difference between hypertensive and normal dogs insofar as the quantity of antidiuretic principle in the urine is concerned. During dehydration it is increased in both. Removal of the posterior pituitary does not interfere with the development of, or lower, experimental renal hypertension in the rat (286). The same holds true for the removal of the entire pituitary in dogs (148, 315). Thyroidectomy, gonadectomy and pancreatectomy do not affect experimental renal hypertension (148).

Of the endocrine organs, the only one that may possibly play a significant, even if only a secondary part, is the adrenal (79), although this conclusion is contested (57, 351) on the basis of inadequate evidence. There is no evidence to indicate that the medulla of the adrenals plays a part (152), but there is definite indication that the adrenal cortex may play a secondary part (148, 431) in experimental renal hypertension, although there are those (350, 351) who do not concede this. Excision of both adrenals in dogs interferes with the development of hypertension due to constriction of the main renal arteries, unless adequate supportive and substitution therapy are given (148). The part possibly played by the adrenal cortex will be discussed further under the heading of the humoral mechanism.

Recently Victor (406) reported the production of hypertension in dogs by ligation of "hilar artery and vein" and of "grossly visible arteries and veins at either the superior or inferior pole" of one adrenal. Although he does not state this, yet it is evident that nerve fibers must have been included in these ligatures. Victor does not explain or speculate about the mechanism of the hypertension. We have tied off the hilar part of the adrenal vein and the vessels going to and from the upper or the lower pole of one adrenal, in seven dogs, without noting any hypertensive effect. This contribution by Victor must await confirmation and elucidation.

The Humoral Mechanism of Experimental Renal Hypertension. The basic observation that constriction of the blood supply to the kidney produces experimental renal hypertension, and the elimination of a possible primary part played by the nervous and endocrine systems, stimulated the search for the humoral mechanism of renal origin most probably responsible for the elevation of blood pressure in this type of hypertension, and possibly also in human essential hypertension. Most recent contributions to the subject have dealt with the nature of this mechanism, about which two distinct viewpoints have been adopted and investigated: 1, that a kidney the seat of certain pathologic conditions may be the source of a substance which, when it enters the circulation, raises the blood pressure (297); 2, that the normal kidney may be the source of a substance the absence, destruction or neutralization of which results in hypertension (172). Most of the evidence, both old and recent, has been in favor of the existence of the first mechanism (252, 378), although there are still some who are opposed to this theory (396).

It has been shown that interference with the blood supply to any other organ but the kidney does not result in either temporary (253) or permanent (152) elevation of blood pressure. Constriction of the celiac axis and superior mesenteric artery (256), of the femoral and splenic arteries (152) and of the aorta below the origin of both main renal arteries (148) does not produce a rise of blood pressure; but constriction of the aorta immediately above the site of origin of the main renal arteries is followed, in about 24 hours, by the elevation of the blood pressure (146, 353, 388). Steele and Cohn (389) believe that the hypertension observed in human cases of coarctation of the aorta is like the effect of constricting the aorta above the origin of both renal arteries. Anastomosis between renal artery and renal vein does not result in hypertension (238). The removal of both kidneys, although profound azotemia is a consequence, is not followed by the development of hypertension; therefore, renal excretory insufficiency by itself, is not a sufficient condition for its development. This is also well demonstrated by the fact that in acute nephrosis, with uremia, produced by various metallic poisonings, in animals and man, the blood pressure rarely becomes significantly elevated.

The first, but indirect, evidence in favor of the probable existence of a renal humoral mechanism for the elevation of blood pressure in experimental renal hypertension was produced by the occlusion of both main renal veins at the same time that both main renal arteries were constricted sufficiently to

produce hypertension. The blood pressure did not become elevated, although the dogs died in profound uremia (148). The arteriolar lesions of the malignant phase did not develop in these dogs. Shunting of the venous blood of a dog from the only kidney, with renal artery constricted, through the liver, by means of a reversed Eck fistula, had no effect in either preventing the development of hypertension or lowering the blood pressure (50, 405). Any important effect of the liver on the pressor substance of renal origin was thus excluded.

A more direct demonstration of a possible humoral mechanism occurred when it was found that if a kidney was transplanted to the neck (24, 26, 105, 211) or groin (142, 336) of a bilaterally nephrectomized dog or rabbit the usual pressor effect resulted when the renal artery of the transplanted kidney, with no nervous connection with the rest of the body, was constricted. Also, the transplantation of a partially or completely ischemic kidney from one dog to the neck of a bilaterally nephrectomized dog, resulted in a temporary elevation of the blood pressure, after the circulation of the ischemic kidney was restored (33).

Corrigan and Pines (64) reported a fall of blood pressure in dogs with experimental renal hypertension, as a result of renopexy (elevation of the kidney for a distance of about 5 cm.) which was supposed to increase the outflow of blood from the renal vein. Why the blood pressure should fall as a result of this procedure is not clear. This experiment requires confirmation before it can be accepted and evaluated. It may have some significance in connection with the problem of hypertension associated with ptosis of the kidney (266). The study of Gabriele (137), however, indicates that such a change of position of the kidney is likely to reduce profoundly, rather than increase, outflow of blood from the renal vein, because the piezometric angle of the renal artery is made relatively acute by this change of position.

It has been shown that when the main renal artery (54, 55, 398), or the entire renal pedicle (120, 136, 248, 253) is occluded, for varying periods, in the cat, rat and dog, a prompt rise of blood pressure follows the release of the clamp and restoration of the circulation through the kidney, and this occurs even in animals with kidney denervated and central nervous system ablated (54). This effect is not prevented by 933F[1] (248) or by sodium cyanide (263); but previous establishment of tachyphylaxis in cats interferes with the phenomenon (248) and it does not occur when similar ischemia of other organs is produced (253). This study was carried further by the demonstration of Prinzmetal, Lewis and Leo (335) and others (233) that saline perfusate of such a totally ischemic kidney produced a definite rise of blood pressure when it was injected intravenously into other animals of the same or different species. Boylston and collaborators (28) did not find a pressor substance in the saline perfusate of a partially ischemic kidney, but this may be explained by the prolonged perfusion which they practiced.

The presence of a vasoconstrictor substance in the blood plasma from the renal vein of a dog with experimental renal hypertension due to incomplete constriction of the main renal artery has been demonstrated by some (32, 105, 197,

[1] This is piperidomethyl-benzodioxane, which is usually referred to as Fournier 933 or 933F.

212, 213 398) and not by others (265), by the Läwen-Trendelenburg technique, in the South American toad. Some investigators (126, 198, 223, 334) failed to demonstrate pressor substance in the systemic blood of hypertensive dogs or human beings and there was also no effect of blood from hypertensive dogs on the tonus of surviving arterial rings (422, 423). But Solandt and collaborators (385) did observe a definite rise in the blood pressure of a bilaterally nephrectomized dog to which was given a direct transfusion from a hypertensive dog; and Braun Menendez and Fasciolo (32) obtained a pressor effect in a normal dog as a result of the intravenous injection of 100 cc. of renal venous blood from the transplanted renal ischemic kidney of another dog.

That the hypothetical effective vasoconstrictor substance is not sympathicomimetic is shown by the fact that 933F (piperidomethyl-benzodioxane) does not have any greater effect on the blood pressure of a hypertensive animal than of a normal one, whereas, in both, the effect of epinephrine is completely reversed by an injection of 933F (119).

All the work outlined above pointed directly to a chemical agent of renal origin as the probable cause of the hypertension. The search for this agent began with the resurrection of an old observation made by Tigerstedt and Bergman (402) confirmed by others (187, 188, 197), namely, the presence of a substance in the crude saline extract of a normal rabbit kidney, capable of inducing a pressor effect when injected intravenously into a normal rabbit. This substance, which Tigerstedt and Bergman called Re-nin' (re'-nin, in English) is now accepted as the basic substance of the humoral mechanism of experimental renal hypertension. Prinzmetal and collaborators (333) and others (433) found a greater amount of this pressor substance in the kidneys of dogs, but not rabbits, with experimental renal hypertension, and of hypertensive persons (393). Although this substance seemed to possess at least some of the basic requirements for the hypothetical pressor substance of experimental renal hypertension, yet its lack of vasoconstrictor properties (121) led two independent groups of investigators (36, 233, 234, 277, 278, 287, 308) to the realization that renin is not a direct pressor substance, but a key substance (241, 245) of the humoral mechanism which acts upon a substrate in the blood, to form a new substance, that is vasoconstrictor and therefore pressor.

In 1940, Landis (236) wrote, "The evidence that renal ischemia raises the blood pressure by a humoral mechanism seems unassailable". Yet Dock (88, 89), as the result of some acute experiments, has arrived at the conclusion that there is no peripheral vasoconstrictor in the blood of the hypertensive rabbit. Because the pithing of a normal rabbit, with blood pressure kept elevated by an infusion of renin, did not result in the fall of blood pressure (unless the infusion was stopped) while it did fall when a hypertensive rabbit (not receiving renin) was pithed, he concluded that in the latter the normal vasomotor mechanism was called into play and the vasoconstriction was abolished. A possible explanation, which has not been tested, is that the immediate effect of pithing is, temporarily at least, to abolish the output of the primary constituent of the humoral mechanism. This should be investigated.

In recent years, as a result of work on shock, evidence has been accumulating

which indicates that the normal kidney may play a part, even in the homeostatic regulation of normal blood pressure, through the humoral mechanism of renal origin which begins with the secretion of renin (184, 207, 215, 284, 355). There are those who doubt this. The low blood pressure in shock (hemorrhagic or otherwise) induces the secretion of renin and the formation of hypertensin (56, 83) in the blood. Shock due to hemorrhage produces a lowering of hypertensinogen (56, 83) in the blood of otherwise normal dogs but not in nephrectomized dogs or in dogs treated with KCN.

A generally accepted terminology for the constituents of the humoral mechanism has not yet been established, and none of the substances has been isolated in pure form. At the present time, the mechanism is supposed to be as follows, and the following names are in use by different investigators for the various substances which constitute this mechanism:

Renin
an enzyme from the kidney, enters the bloodstream through the renal vein, and acts upon

Hypertensinogen
Prehypertensin,
Hypertensin-precusor
Renin-substrate

a globulin in the blood plasma, to form

Hypertensin
Angiotonin

a polypeptide, which is the active vasoconstrictor substance, and which can be inactivated by

Hypertensinase
an enzyme in the blood and in extracts of some organs.

The name renin activator, originally proposed for the stubstrate, by Page, should certainly be abandoned, because the term is misleading, since the substance it represents does not activate anything, and is actually acted upon by the renin. It has been shown that it is contained in the α_2 globulin fraction (327).

Agreement upon a single nomenclature would be highly desirable, and the author suggests the acceptance of renin, hypertensinogen, hypertensin and hypertensinase as the most appropriate names for the constituents of the humoral mechanism of renal hypertension.

The following is a brief summary of the chemical and physiological properties of these substances:

Renin. Chemical properties. Renin has not yet been isolated in pure form, although several attempts to purify it have been made (22, 52, 84, 104, 191, 228, 322, 358). All that can be said is that it is either a protein, a protein-like substance, or contains protein as a contaminant (181, 191, 252, 321). It is present in the press-juice (187, 188) or in various extracts (35, 196, 218, 241, 321, 325, 329, 396, 441) of renal cortex (not medulla) and does not occur in similar extracts of other organs (spleen, liver) (121). It is thermolabile (destroyed by heat above 56°C) and not dialyzable. In an electric field, renin migrates to the cathode, the isoelectric point being between pH 6.5 and 7.5. The results of studies on the behavior of renin subjected to precipitation with ammonium

sulphate and dialysis suggest that it is a pseudoglobulin (181) but this has not been established beyond question. Most investigators regard it as a proteolytic enzyme (278, 309) extractable by various methods from all mammalian normal kidneys and from the kidneys of other animals (321, 434, 440). Plentl and Page (328) found in their purified renin the enzymes pepsin, trypsin, cathepsin, aminopeptidase and carboxypeptidase. They were not able to identify one of these enzymes as being specifically the enzyme renin, although they were able to exclude aminopeptidase and carboxypeptidase. But Winternitz and collaborators (441) have asserted that renin, in very pure form, is free of ordinary enzymatic activity. Various methods for extraction and purification have been described (52, 181, 252).

The preparation and isolation of renin, by extraction, are best carried out under aerobic conditions, because anaerobic conditions favor the formation of pressor amines, tyramine, isoamylamine and probably ethylamine (91). Kidneys from young rats have been reported to contain more renin than those from old rats; but old rats are more sensitive to parenterally administered renin (182). This study should be repeated on rats and other animals by use of more recently developed methods for extraction and assay.

Physiological properties. Renin is effective in raising blood pressure, when it is injected intravenously into animals, but it is not vasoconstrictor when dissolved in saline and perfused through a circulatory system from which the blood has been washed out. It is not a direct pressor substance (36). The physiological effects of renin injected intravenously into a normal animal are identical with the hemodynamic state of animals with experimental renal hypertension. It is not species specific for animals; all renin, including human, when injected intravenously, produces a pressor effect in all animals. But no animal renin so far produced has a pressor effect in man, who responds only to human renin (403). The pressor effect of renin is produced indirectly by the interaction of renin and hypertensinogen, a globulin in the blood plasma. This results in the formation of a new substance, hypertensin, which is the effective vasoconstrictor and therefore the true pressor substance. The site of action of renin itself is therefore in the blood and not on the vasomotor endings of the peripheral arterioles (402) or directly on the arteriolar muscle (2, 295), as was thought at first. In the rabbit, the pressor effect of renin is abolished or lowered by ether anesthesia (357), but this is not true of the dog.

Tigerstedt and Bergman (402) reported that renin exhibits tachyphylaxis, which means a decrease of the response of the animal to repeated intravenous injections of the same amount of renin, provided these are made only a few minutes apart. This has been confirmed by others (191, 197, 228, 240, 241, 268, 295). This effect is independent of anesthesia, type of anesthetic, type and method of preparation of renin and occurs in a bilaterally nephrectomized animal (413). It produces no direct effect on the isolated perfused heart (197, 200) but when injected intravenously into an animal, increases the force of cardiac action. When it is injected intravenously, it produces its effect even in an animal with spinal cord sectioned or destroyed (272), and its pressor effect

is not the result of the stimulation of an endocrine organ. It produces its effect in the absence of hypophysis and adrenals (272). In the rat, it is said to be effective when injected intraperitoneally, and more effective in the bilaterally nephrectomized than in the normal rat (115, 272). This should be tested on dogs and other animals. It does not produce a reduction in peripheral blood flow or a fall in the temperature of the surface of the skin (221). Its effect is not reversed by 933F (222), or ergotamine tartrate, nor is it enhanced by cocaine. It has a more prolonged vasopressor effect in bilaterally nephrectomized animals (348), probably on account of the increase of hypertensinogen in the blood of such animals. Renin injected intravenously into the normal rabbit produces diuresis and increased excretion of sodium chloride (323). It has been asserted that extracts containing renin from the kidneys of animals with experimental renal hypertension and of acutely and completely ischemic kidneys contain more renin than the extract of the opposite normal kidney, or of normal kidneys from other animals. But Landis (237) was unable to correlate the renin content of the kidney and the physiological state of the kidney. This work should be repeated with the use of the more recently devised methods of extraction and assay.

The presence of renin in the renal venous blood and systemic blood of animals with experimental renal hypertension has been inferred from the production of hypertensin, when the blood was added to hypertensinogen and incubated with it (80, 105, 298). Dexter and Haynes (82), by this indirect method, measured. the amount of renin in the systemic blood of human hypertensive patients. They found renin only when the blood pressure was rising abruptly (eclampsia and acute glomerulonephritis); none in the blood of patients with benign essential hypertension. On this account they suggested that renin may be involved only initially, during the development of hypertension. The existence of renin in renal venous blood from a human kidney, the renal vein of which had been occluded for 12 minutes, has been reported (338). The criticism of most of the negative results is that too much renin is expected and therefore too small a quantity of blood has been tested. Gollan, Richardson and the author (unpublished), by incubation of larger quantities of blood, and with proper precautions to prevent the action of hypertensinase, have been able to demonstrate the formation of hypertensin by the action of renin on the natural hypertensinogen in the blood of dogs with experimental renal hypertension, both benign and malignant. In the animals, in the malignant phase, there was much more hypertensin in the final product, and therefore presumably more renin in the blood. A more direct way of determining the renin content would be to measure the rate of formation of hypertensin.

The action of renin is not affected by hypophysectomy, thyroidectomy, splanchnicectomy, gonadectomy, splenectomy, pancreatectomy, abdominal evisceration or destruction of the spinal cord (272, 278, 435). Bilateral adrenalectomy, although it has no immediate effect (272), is followed by a progressive decrease in the response to the intravenous injection of renin (210, 278). By slow intravenous administration of a small amount of renin, the blood pressure

was kept elevated about 30 mm. above normal, for four hours. Larger quantities produced untoward circulatory disturbance (201). With one exception (430), investigators agree that the effect of the purest renin so far produced is not potentiated by cocaine or reversed by ergotamine tartrate or 933F. Its ultimate effect is therefore different from that of known sympathicomimetic pressor amines. Its action differs from that of adrenalin, pituitrin and tyramine. The pressor response to an intravenous injection is more delayed because formation of the true vasoconstrictor substance, hypertensin, has to occur during this interval. The rise of blood pressure is not as steep and lasts much longer. It takes about two minutes for a unit of renin to produce its maximum rise of 30 mm. Hg direct mean pressure and the gradual return to normal may take as long as 30 minutes. The rise of blood pressure is not associated with a decrease in peripheral blood flow.

The action of the purest renin so far produced (191, 197, 228, 240, 241, 268, 395, 435, 441) is no different from that of crude renal extracts that do not also contain a depressor substance. According to Grollman and collaborators (172, 433) the amount of renin in a kidney extract, or its apparent concentration, depends upon the care of the tissue prior to its preparation, the method of preparation and the duration of time between the preparation of the extract and its assay on a test animal. The last part of the statement must apply only to very crude extracts, because the author has assayed moderately purified liquid, frozen or lyophilized preparations of renin several times during a period of three years without noting much deterioration in any of them. All preparations of renin were sterilized originally by filtration through a Seitz filter, and kept in sterile condition, without the addition of preservatives, in sealed vials. The liquid and lyophilized renins were kept at 4°C., while the frozen renin was kept at −30°C.

Fate of Renin in the Body. The finding of Tigerstedt and Bergman (402) that bilateral nephrectomy increases the sensitivity of the organism to an intravenous injection of renin has been confirmed by others (115, 272, 310, 411); but their idea that the kidney eliminates renin has not been confirmed, because it has been found that only when large quantities of renin are suddenly injected intravenously does it appear in small quantity in the urine. The maximum excretion occurs in about one hour after the injection (209). Because the absence of the liver and kidneys delays the rate of disappearance of renin injected intravenously, Leloir and collaborators (247) concluded that both play an important part in the elimination of renin from the body. The main mechanism of elimination of renin from the circulation is now considered to be destruction by the tissues of the body, including the kidney and liver (209).

The Unit of Renin. Various units have been suggested and different methods for the purification and bioassay of animal renin have been described (52, 151, 181, 191, 228, 239, 246, 267, 272, 321, 324, 329, 343, 359, 395, 410). A special method for the assay of human renin has been described (84, 280). It would be highly desirable to determine the best method, and agree upon a unit, for animal and human renin, in order to have at least one good basis for the comparison of

results obtained by different investigators. Leloir and collaborators (246) defined a unit of renin as the quantity which, when incubated with hypertensinogen in excess, at 37°C. and pH 7.5, produces 0.5 unit of hypertensin in two hours. We have used the direct method for the assay of animal renin (151) and have defined the dog unit as the quantity which, when injected intravenously into normal unanesthetized dogs, will cause a rise of direct mean pressure of 30 mm. Hg. This has proved a very useful method for purification studies (228). A method for the study of the reaction kinetics of renin has been described by Plentl and Page (325).

Hypertensinogen. This is a protein (327) of the blood plasma upon which renin acts to form the vasoconstrictor substance hypertensin (angiotonin). It is thermolabile, not ultrafiltrable, and not dialyzable. It is a pseudoglobulin (34, 278, 279, 360) and a constituent of the α_2-globulin fraction (327). Serum albumin and hemoglobin do not act as renin substrate. Edman (97), however, considers that the method and conclusions of this study by Plentl, Page and Davis (327) are "not entirely above criticism". A method for its preparation in relatively pure form, from beef serum, has been described (360). It is precipitated from serum by ammonium sulphate between 0.30 and 0.41 saturation, at pH 6.8, and does not precipitate after dialysis against distilled water.

Renin extracted from the kidneys of mammals reacts to form hypertensin with hypertensinogen from any mammal, with the exception of man. Human renin is active on the hypertensinogen of any mammal including man. Renin is found in the kidneys of chickens and ducks and it reacts with hypertensinogen from any of those birds but not with mammalian hypertensinogen. Mammalian renin is not active on avian plasma. There seems to be no renin in the kidneys of toads or sharks (14).

The reaction between renin and hypertensinogen is enzymatic and specific in the sense that only blood plasma or serum of various kinds acts as the substrate. Hemoglobin, liver, spleen, thymus, testis, lung, heart, skeletal muscle and vegetable proteins do not act as renin substrate. It requires only a small amount of renin in proportion to the amount of substrate in the blood to bring about the optimum reaction, and the amount of hypertensin formed is proportional to the amount of blood globulin. The action is not greatly affected by temperatures up to 37°C, for although it is hastened by incubation, at 37°C, yet it is only slightly retarded, and goes on to completion, at 0°C, even in the presence of hypertensinase (356). (See under Hypertensinase.) The effect of renin on hypertensinogen cannot be reproduced by pancreatin, papaine or extract of liver or spleen, nor by pepsin at the pH of the blood (see Pepsitensin).

According to Leloir and collaborators (245), the quantity of hypertensinogen is about the same in the blood of normal dogs, cows and pigs, while the horse has a little less. Lymph contains as much hypertensinogen as plasma (134). The amount of hypertensinogen is increased in the blood of bilaterally nephrectomized animals (245, 278). It is not affected by hypophysectomy. Adrenalectomized rats are less sensitive to an injection of renin (431) and hypertensinogen was found greatly decreased in bilaterally adrenalectomized

animals (138, 148, 245, 252) and in dogs with hemorrhagic shock (56), in human hepatic insufficiency (190), but not in one case of Addison's disease (189). It is not increased in the blood of hypertensive animals or in human hypertension (189). Adequate substitution therapy of adrenalectomized animals with adrenal cortical extract results in the return of hypertensinogen in the blood to the normal level. Repeated injections of renin, with final tachyphylaxis, are said to deplete the hypertensinogen in the blood (278). It is probable that hypertensinogen is formed in the liver (245). Hepatectomy and destruction of the liver by chloroform poisoning cause hypertensinogen to decrease in amount and even to disappear from the blood (148, 252, 314) and it is often significantly decreased in human hepatic insufficiency (189). Hepatectomy causes a decrease, but if bilateral nephrectomy is performed at the same time, then there is no diminution, because no renin is being liberated to use up the hypertensinogen (247).

Unit of Hypertensinogen. This may be defined as the amount of hypertensinogen which, with an excess of renin, and in the absence of hypertensinase, will produce a unit of hypertensin.

Hypertensin (Angiotonin). The substance formed by the action of renin on hypertensinogen is the final effector, vasoconstrictor, and therefore pressor substance of the humoral mechanism of renal hypertension which causes the increased tonus of the smooth muscle of the arterioles when renin is injected into, or presumably enters, the blood stream (279). It was discovered independently by two groups of workers who gave it the names hypertensin (36) and angiotonin (308). If it is finally proved that the product which results from the interaction of renin and the substrate in the plasma of hypertensive animals and man is the substance that causes persistently elevated blood pressure, then the term hypertensin of the South Americans (279) will be the more appropriate. It has already been shown (36) that when hypertensinogen is added to the blood plasma or serum from the blood of a totally ischemic human kidney, and the two are incubated under conditions that prevent the action of hypertensinase, a substance results which has all the properties of hypertensin (338). Evidence is accumulating that renin, (and therefore hypertensin) is present in the systemic blood of hypertensive animals and man. This is one of the most important problems that remain to be settled.

Chemical properties. Hypertensin is soluble in water, 96 per cent ethanol, 75 per cent acetone, liquid phenol, glacial acetic acid and methylene glycol, but insoluble in ethyl ether, chloroform and amyl alcohol. It is dialyzable and thermostable (resistant even to boiling temperature), alkali-labile, (destroyed completely by boiling for 1 hour at pH 10), relatively acid-stable (does not become inactivated by boiling for 2 hours in 1N HCl), can be precipitated with phosphotungstic acid, is destroyed by oxidation and acetylation with ketene (331), but is not affected by reducing agents (309). It is fluorescent and gives the color reaction for arginine, but it is not destroyed by the action of arginase (73). A positive Sakaguchi reaction is its only reaction for protein (309). It contains tyrosine and histidine, but not cysteine, tryptophane or proline (73). In

electrophoresis experiments, hypertensin behaves as an amphoteric electrolyte of neutral character (73). Cruz Coke has outlined the similarities and differences between hypertensin and the pituitary pressor hormone (73). That the reaction between renin and hypertensinogen is disintegrative is indicated by the fact that pepsin, a known proteolytic enzyme, may replace renin in this reaction, but only at a very low range of pH. Pepsin is inactive in the optimum range of pH for the activity of renin, and does not produce a pressor effect upon intravenous injection. That the reaction between renin and hypertensinogen is enzymatic is now generally accepted (325). Hypertensin is also destroyed slowly by hydrolysis, by proteases and peptidases (73). It is now considered to be a polypeptide of low molecular weight (244), although Page and Helmer (309) found that it gave a negative biuret. It is destroyed by pepsin, trypsin, papain (36, 96, 278) and extracts of some organs, as well as by fresh normal blood plasma, serum and laked blood corpuscles, all of which contain hypertensinase. The vasoconstrictor effect of hypertensin is inactivated by amino-oxidase from liver of sepia officinalis and tyrosinase from Psalliota campestris (66). It is easily destroyed during purification, for unknown reasons (244). Cruz Coke (73) lists a number of chemicals which partially or completely inactivate hypertensin, the most important being iodine. The effect of the latter is due to the transformation of the tyrosine in hypertensin. He also emphasizes the accelerated destruction of hypertensin by hypertensinase upon the addition of oxidized cytochrome C to renal extract containing hypertensinase.

Crystallization of angiotonin in the form of its picrate and oxalate has been reported by Page and Helmer (309), but this result was not confirmed (330). Recently a preparation of high potency was obtained by Edman (97). As little as 0.5 γ elicited an appreciable effect on the blood pressure of a cat. Cruz Coke (73), by means of anionic resin has been able to purify hypertensin, so that it had only 0.08 mgm. N per unit. Cationic resin adsorbed the hypertensin, but it was not possible to elute it. The melting point of this substance in pure form has not yet been reported. The effect of proteolytic enzymes on partially purified angiotonin has been investigated (330), but the authors state that "such a study cannot be taken as offering proof of the structure of this substance".

The most recent important contribution to the subject of the purification of hypertensin is by Edman (97) whose method consists of chromatography, precipitation with nitranilic acid and electrodialysis. He obtained in this way a 600–700 fold purification of the active principle, but the yield was only 3 per cent. This great loss he attributes to the lability of the hypertensin structure. His highly purified hypertensin is almost 40 times more effective in raising blood pressure than a comparable amount of tyramine phosphate. He regards his product as essentially pure, judging by its behavior on partition chromatography on filter paper. By means of electrophoresis he determined the isoelectric point to be at pH 6.8. From the diffusion constant of his hypertensin in pure water he calculated the molecular weight to be 2700. By means of paper chromatography, and some preliminary quantitative determinations, he found

in his hypertensin the amino acids lysine, histidine (28 per cent), glycine, alanine (3.8 per cent), serine, proline (5 per cent), valine, tyrosine (2 per cent), leucine (or isoleucine), aspartic acid (4.5 per cent) and glutamic acid (5 per cent).

Physiological properties of hypertensin. The pressor effect produced by an intravenous injection of hypertensin is immediate and steep, like that of adrenalin. The maximum rise from the intravenous injection of a unit of hypertensin occurs in a minute or less and the return to normal in three minutes or less. Accompanying the hypertensive effect of hypertensin, as in the case of renin, there is no change in blood flow through the skin, as measured by temperature of the skin and by direct observation of blood vessels growing in a wound in the skin of a rabbit's ear (301). There is a decrease in the volume of the spleen and kidney (36), a decrease in coronary blood flow, increase of venous pressure and decrease of renal blood flow. Hypertensin produces no significant symptoms, except forceful heart beat, when it is injected intravenously into man (301). In vagotomized dogs the heart rate is not modified (36). Continuous intravenous injection produces an increase of blood pressure which is maintained during the period of injection (36). On isolated intestinal strips hypertensin has a stimulating effect(309) and a similar effect is produced by it on nearly all smooth muscle.

Hypertensin shows no specificity. This product of the reaction between mammalian renin and hypertensinogen shows its characteristic pressor effect in dogs, chickens, toads and snakes. Dogs and chickens do not exhibit tachyphylaxis as a result of repeated injections of hypertensin. Toads and snakes are rapidly rendered tachyphylactic (14).

The assertion of Page and Helmer (301) that repeated injections of angiotonin produce tachyphylaxis by the method of perfusion of the rabbit's ear has not been substantiated by the work of others (278) with the repeated injection *in vivo*. The method of perfusion of the rabbit's ear is notably unreliable and of questionable value in this connection. Existence of angiotonin activator, postulated by Page (301), was not confirmed. As in the case of renin, injected intravenously, the pressor effect of hypertensin is unaffected by cocaine (191,309), atropine (309), ergotamine (191), 933F, or stilbestrol (309). It is therefore not a sympathicomimetic substance. The effect is greatly increased in bilaterally nephrectomized dogs. Hypertensin, like renin, when injected intravenously, induces a pronounced but transient rise in potassium and a greater and more prolonged rise of sugar in the blood of dogs under the influence of chloralose. No special significance has been attached to these observations and they have not been confirmed (113). The striking difference between renin and hypertensin is that hypertensin produces vasoconstriction when it is added to Ringer's solution and perfused through an isolated organ from which the blood has been washed out, while renin in Ringer's does not have this effect. The direct effect of hypertensin and the indirect effect of renin are exerted on the musculature of the peripheral blood vessels. Large doses of hypertensin injected intravenously into normal persons caused a decrease in blood volume and cardiac output, as measured by the ballisto-cardiograph (400). The effect

of hypertensin injected intravenously is not affected by vagotomy, excision of the carotid bodies, splanchnic nerves, pituitary, pancreas, liver or adrenals, by evisceration, destruction of the adrenal medulla or rapid elimination of the central nervous system, in cats and dogs (97, 307).

The finding of Page and collaborators (233, 295, 298, 299, 300, 302, 304) that there is in the peripheral systemic blood of hypertensive animals and man a vasoconstrictor "angiotonin-like substance", not present in the blood of normotensives, has not been confirmed. There is no proof that this hypothetical substance is identical with renin or hypertensin. These studies were also carried out by the rabbit's ear perfusion method, so the results will require confirmation by some other method before they can be accepted. The same holds true for the demonstration of a pressor substance in blood perfused through an isolated kidney with renal artery constricted, demonstrated by the same method (233).

Gregory and collaborators (162) have failed to demonstrate an increase of vasoconstrictor substance for frogs in the ultrafiltrate of blood plasma of human beings with essential hypertension of long duration. Basing themselves on the results of a quantitative method of questionable value, they concluded, without full justification, that essential hypertension is probably not caused by an increased production of hypertensin. Dexter and Haynes (82) failed to produce hypertensin when they added hypertensinogen to the systemic blood of patients with essential hypertension. For the present, at least, the most obvious criticism of this work is that the amount of blood tested has always been too small. More studies of larger amounts of blood are necessary. The failure of Grollman and Rule (177) to demonstrate a pressor effect in the normal rat of parabiotic twins, of which one was hypertensive, on a renal basis, does not disprove the existence of a humoral mechanism in this type of hypertension and does not prove their contention that the normal kidney elaborates a substance the absence of which results in hypertension. The explanation of their results may be simply the slow passage of the hypertensin from one animal to the other and the rapid destruction of that which does pass across (31).

Unit of Hypertensin. This is best defined as the quantity of hypertensin which, when injected intravenously in the unanesthetized trained dog, will give a rise of 30 mm. Hg direct mean femoral blood pressure. Braun-Menendez and collaborators (36) consider an elevation of 20–30 mm. as a unit, because of a greater variability of response to the same dose than we have noted. This difference may be due to the fact that our animals are well trained and lie quietly during a determination. This is a necessary condition for uniform results from injections of renin or hypertensin. They have published a curve of the pressor effects of various amounts of hypertensin from which the number of units is estimated by interpolation (36). This is not accurate, in our experience, since the pressor response to hypertensin is not a linear function of the concentration of hypertensin. It would be highly desirable to agree upon one method.

Pepsitensin. It has been shown by Croxatto and collaborators (68) that pepsin, if it acts on blood globulin, at pH 3.0, or lower, can produce, by peptic

digestion of the protein, a vasoconstrictor and pressor substance which they have called pepsitensin and which they assert is identical in chemical and pharmacological properties with hypertensin, except that pepsitensin is not inactivated as rapidly by hypertensinase (6, 37). This difference is attributed to impurities in pepsitensin and it is asserted that the purer the pepsitensin the more sensitive it is to inactivation by hypertensinase and other proteolytic enzymes (73). Hypertensin is more easily destroyed than pepsitensin by the action of pepsin. This is to be expected, since pepsitensin results from the action of pepsin on blood globulin. Renin and pepsin, however, do not act on exactly the same substrates. Plasma globulins acidified to pH 2 and then neutralized, or precipitated with alcohol and redissolved, no longer yield hypertensin with renin, but still yield pepsitensin with pepsin. It would seem that pepsin acting on denatured globulin is capable of producing a pressor substance, while renin fails to do it (244). The existence of pepsitensin has been confirmed by others (192, 424), including ourselves. What significance these experiments will have eventually in connection with the humoral mechanism of renal hypertension it is difficult to estimate at this time. For the present they merely indicate the probable non-specific character of the renin-hypertensinogen reaction, and the disintegrative nature of this reaction, which results in the formation of an active pressor substance.

Hypertensinase. It was noted by Page and Helmer (310) that when renin was incubated with plasma or serum, angiotonin was formed, but that continued incubation resulted in the destruction of the angiotonin. They attributed this to the effect of the renin. Muñoz and collaborators (278), however, showed that this inactivating effect of the renin could be eliminated while the capacity of the renin to produce the pressor substance remained unaffected. This led them to postulate the existence of another enzyme associated with impure renin. For this enzyme the name hypertensinase was chosen by the South Americans (107), and it has been called angiotonase, to correspond with angiotonin, by Page and collaborators (313).

Hypertensinase is a hydrolyzing enzyme, or group of enzymes (107, 278), present in blood plasma and serum, in laked red blood corpuscles, and in extract of organs, especially intestine, kidney, pancreas, spleen and liver, with the ability to destroy hypertensin *in vitro*. Intestinal mucosa is the richest source of this enzyme, while blood serum and plasma that are not hemolyzed contain only a relatively small amount (107). There is no increase of hypertensinase in hypertensive dogs (81) or man (190).

Properties. Hypertensinase is a protein and an enzyme, or mixture of enzymes, not dialyzable, and precipitable by half saturation with ammonium sulphate. It is acid and heat labile (destroyed at 60°C. or over). The activity of hypertensinase is optimum at a pH between 7.5 and 8.5. This enzyme is inactivated, if kept at pH 3.9 for 15–20 minutes, at 37°C., even in the presence of renin and hypertensinogen. The action of hypertensinase is greatly decreased at 18°C. (13) and becomes negligible at 0°C. (356), at which temperature the reaction between renin and hypertensinogen is only slightly retarded, but may

go on to completion. Although hypertensinase has not yet been isolated in pure form, Croxatto and collaborators (69, 70, 71) have concluded that its effect on hypertensin is due to a peptidase, and most probably an aminopeptidase action (72). This agrees with the finding that the hypertensin inactivating effect of an extract of pancreas is due to carboxypeptidase (330). Their conclusion, on the basis of inhibition experiments, is that the hypertensinase effect of renal extract and of laked red blood corpuscles is partly proteolytic and partly an oxidase effect. From renal extracts it is difficult to obtain hypertensinase free of renin, but the hypertensinase may be easily destroyed in a solution of renin. There is some probability that the kidney may be the main source of the hypertensinase in normal blood plasma. This is indicated by the fact that there is almost complete absence of hypertensinase in the blood of bilaterally nephrectomized dogs. There is less hypertensinase in lymph than in plasma (134). The fact that the blood from the ischemic kidney contains less hypertensinase than normal blood may mean that the ischemic kidney produces less (426). The exact part, if any, played by hypertensinase in the maintenance of normal or elevated blood pressure is not yet elucidated.

The Unit of Hypertensinase. This has been defined as the smallest amount of hypertensinase which is capable of inactivating one unit of hypertensin in 30 minutes at 37.5°C. The technique for the determination of this unit is to combine variable amounts of hypertensinase with a constant amount of solution known to contain one unit of hypertensin. The mixture is incubated for thirty minutes at 37–38°C. and is then injected intravenously into a normal, unanesthetized trained dog. The least amount of hypertensinase which destroys the entire unit of hypertensin, so that there is no rise of blood pressure after its intravenous injection into an animal, is the unit of hypertensinase (107).

Mechanism and Site of Formation or Release of Renin. Despite the vast amount of work that has been done on the properties of the various constituents of the humoral mechanism of experimental renal hypertension, but little of a definite nature has been established beyond question about the mechanism and site of formation and release of renin, upon the action of which the formation of hypertensin depends. Although Leloir (244) states unequivocally that most of the known facts about the mechanism of action of renin are consistent with its being the cause of renal hypertension, yet Grollman (166) even questions the existence of preformed renin in the kidney, and considers that it is merely the product of autolysis *in vitro*. His experiments may merely indicate that in the living animal a renin precursor, *prorenin,* as it leaves the living cell, undergoes a change which transforms it into renin, and that the same transformation can also occur *in vitro*. Analogous transformations are known to occur in proteolytic enzymes; for example, the formation of trypsin or pepsin from their inactive precursors.

Exactly what takes place in the kidney, with main renal artery constricted adequately to give hypertension, without causing obvious degeneration or necrosis in the substance of the kidney, is not well known. The observation of decreased oxygen consumption by the ischemic kidney (249) or by ischemic renal tissue (140) has been confirmed (339); but the result has been questioned

on the ground that the reduction may be due to death of a certain number of cells and not to interference with the function of all (264). Inhalation of 7–10 per cent O_2 did not cause a greater rise of the blood pressure of hypertensive dogs (390), and continuous inhalation of 100 per cent oxygen for 48 hours failed to lower their blood pressure. This has been interpreted as not being in favor of a hypothetical anoxemic factor in the pathogenesis of experimental renal hypertension. Yet Cruz Coke (73) has arrived at the view that tissue anoxia, especially renal, plays an important part in the humoral mechanism of renal hypertension. The criticism of the experiments on oxygen consumption may also possibly apply to the demonstration of a great diminution of the cytochrome C concentration and the activities of the cytochrome oxidase and succinic dehydrogenase in slices and homogenates of the kidneys of hypertensive dogs (340). More work of this kind must be done before its significance can be evaluated.

The site of origin of renin, at least *in vitro*, has been investigated, and the evidence at the present time is that it originates in the cortex of the kidney and especially in the lining epithelium of the convoluted tubules. The fact that extract of aglomerular midshipman fish kidney contains no renin (130) has provided no clue to the origin of renin because it has been found that marine fish kidneys which do possess glomeruli do not contain renin, while the kidneys of fresh water fish possess it in great abundance. The explanation for this difference has not been found (132). Yet renin can be produced from the kidney of the dolphin, a marine mammal (98). Because the renin content of the involuting tubular portion of the mesonephros of pig embryo decreased, while that of the developing tubular portion of the metanephros increased, the conclusion was reached that the convoluted tubules are most probably the site of origin (production and storage) of renin (220). In keeping with this is the failure to extract renin from kidneys in which the proximal convoluted tubules had been destroyed by sodium tartrate poisoning (131).

The nature of the stimulus for the release of renin has not yet been determined. The idea that reduction of intrarenal pulse pressure, rather than decreased blood flow to the kidney, is what determines the release of renin and the formation of the vasoconstrictor substance, depends entirely upon the demonstration of a pressor substance in the blood by the rabbit's ear perfusion method; and since, by admission, this is not really a test for renin or angiotonin, it is questionable what significance can be attached to these experiments (62). The assumption of a presumable change from intermittent to continuous pressure beyond the site of the afferent preglomerular arterioles is not justified, for the very reason that a pulsatile pressure in the glomerulus has never been proved to exist. Braun-Menendez (31) states unequivocally that, "The idea that diminished pulse pressure within the kidney causes the liberation of renin has no solid experimental proof". The reduction of blood flow through the functioning components of the kidney (glomerular and peritubular capillaries) is another possible stimulus (152, 250). That there is a reduction in the blood flow through the kidney in most cases of essential hypertension, affecting both kidneys equally (47, 48, 63, 155, 364), as well as in the early stage of experimental renal hypertension (152),

is an established fact, but there is still a question about whether permanent reduction of the blood flow is necessary for the persistence of hypertension in animals (369). This must await better and more direct methods than are available at present to allow frequent determinations of renal blood flow before and after constriction of the renal arteries. Because, in an occasional animal, there was no permanent reduction in the blood flow through the kidney, although there was still a slight increase in blood pressure, some authors (61, 62) have brushed aside the idea that some reduction of blood flow is a necessary condition for the development of the hypertension. For the demonstration of true renal ischemia Chasis and Redish (48) require that the ratio of renal plasma flow (diodrast clearance) to tubular excretory mass (maximum tubular secretion of diodrast) should be calculated, because the reduction of diodrast clearance does not necessarily mean renal ischemia. Smith and collaborators (153, 155, 383) state that the evidence favors the view that the renal ischemia, so frequently observed in essential human hypertension, is a secondary event, and that the primary event is the circulation of a humoral substance of unknown origin which brings about the efferent arteriolar spasm and progressive and parallel reduction in renal blood flow, which they consider characteristic of essential hypertension Others believe that the efferent arteriolar spasm is due to the angiotonin produced by renin in the blood (193), but in experimental renal hypertension this begins only after the renal artery is constricted, and therefore, in man, it should begin only when the renal arteries and afferent arterioles are sufficiently diseased to reduce the size of the lumen.

The demonstration (65) *in vitro* of a normally perfusable vascular bed in the kidneys of human beings with benign hypertension, especially when the perfusing fluid is kerosene, certainly does not justify the conclusion that the perfusion of blood through the kidney *in vivo* is also normal in such individuals. It is interesting, however, that despite the obvious objection to the method, there was a great decrease of the rate of perfusion, even of kerosene, through the kidneys of patients with uremia, due to arteriolosclerosis, glomerulonephritis or pyelonephritis.

The problem of the possible part played by the juxtaglomerular apparatus in the humoral mechanism of experimental renal (156, 157) and human hypertension (229) is by no means settled. This apparatus has been described in detail by Goormaghtigh (156, 158 158a) and others (94, 95, 269).

An increase in the size of the "polkissen", or juxtaglomerular apparatus, and in the number and size of the afibrillar, and sometimes granular, or vacuolated, cells in this apparatus, in the kidneys of dogs with renal hypertension, produced by constriction of the main renal arteries, has been described (158). Goormaghtigh (158a) considers that these cells may have a local, or even general, secretory, or humoral activity and may therefore have a direct relationship to the hypertensive principle. In the rabbit, in which afibrillar cells are common in the normal kidney, Dunihue (95) states that Goormaghtigh has found an increase in the number of afibrillar and granular cells in the juxtaglomerular apparatus of rabbits made hypertensive by constriction of the main renal arteries. Goormagh-

tigh considers that the afibrillar cells have to do with arteriolar tone and that the granular cells are a source of the pressor substance. These findings in the hypertensive rabbit have been confirmed by Dunihue (94, 95) who subscribes to the same views. Kaufman (229), on the basis of an anatomical study of kidneys from normal and hypertensive persons, reached a similar conclusion. Graef and Smith (160, 382), however, have drawn attention to the great variation in the appearance of the arteriolar media and the size and structure of the juxtaglomerular apparatus in normal man and animals and have cautioned that, because of species differences, e.g., the absence of granular cells in the kidney of man and dog, the interpretation of the ischemic changes must be contingent upon a more complete study of normal kidneys. The development of cytologic changes in the juxtaglomerular apparatus, interpreted by some investigators as indicative of endocrine activity, does not constitute convincing proof of the origin of an endocrinogenic pressor substance or precursor in this structure.

There is certainly no direct, convincing evidence that any special cells in the juxtaglomerular apparatus or preglomerular arterioles are the source of a chemical factor which constricts the afferent or efferent arterioles, thus regulating glomerular blood flow, and, presumably, also the peripheral arterioles, with resultant systemic hypertension. As a matter of fact, the presence of renin in the developing pig embryo, in which a definite juxtaglomerular apparatus has not been identified, militates against this view (220). More investigation is is necessary before the functional significance of the juxtaglomerular apparatus can be properly assessed.

The Treatment of Experimental Renal Hypertension by Renal Extracts. After the production of experimental renal hypertension was accomplished, it was hoped that methods of treatment which would be successful for hypertension in animals might be applied with equal success to human hypertension. It was soon shown that all the methods which had previously failed to affect human hypertension were equally unsuccessful in experimental renal hypertension (150, 173), but that the difficulties involved in the clinical appreciation of the hypotensive properties of various agents in essential hypertension are greater than in experimental renal hypertension (1).

Any attempt to treat experimental renal hypertension by affecting the humoral mechanism should take into consideration the following possibilities outlined by Muñoz and collaborators (278): 1, suppression, diminution or inactivation of renin; 2, inhibition of the reaction between renin and hypertensinogen; 3, diminution of the amount of hypertensinogen; 4, inhibition of the action, or destruction, of hypertensin, by an increase in the amount or activity of hypertensinase, or some other agent capable of accomplishing this. Methods 1 and 4 have received greatest attention. The fourth method will be discussed first.

Extracts of kidney had been used empirically, with variable results, for the treatment of human hypertension (150) a long time before persistent hypertension was produced in animals. The first attempts to treat experimental renal hypertension by means of renal (and muscle) extracts were made by Harrison and collaborators (174, 175, 176, 179, 180, 185) who reported a lowering

of blood pressure in hypertensive rats by renal extract given by mouth, or parenterally. Their attempt was based mainly on the assumption (179) that in the temporary hypertension due to unilateral constriction of the main renal artery, the normal kidney, by some humoral mechanism, e.g., by inhibiting or destroying the pressor substance, might be playing a part in eliminating the hypertensive effect of the kidney with renal artery constricted. Other evidence for this view is as follows: the removal of a normal kidney in the presence of one with renal artery constricted causes the blood pressure to remain permanently elevated; after the removal of the ischemic kidney, if the other is normal, the blood pressure falls to normal in six hours, but it takes five times as long for the blood pressure to reach normal after the removal of both kidneys, when one or both are ischemic. This led Katz and collaborators (223, 226, 349) and others (106, 404, 405) to the idea that by counteracting or neutralizing the pressor substance the normal kidney might play a part in the elimination of the chemical mediator of experimental renal hypertension. They believed that the more rapid fall of the blood pressure to the normal level, in the presence of normal kidney tissue, as contrasted to the slow fall in its absence, might be due to the excretion, or destruction *in vivo*, of the pressor substance, by the remaining normal kidney. They showed that in an animal with a ureteral-venous fistula on the side of the normal kidney, the removal of the opposite ischemic kidney was still followed by rapid (6 hours) fall of the blood pressure to normal. They concluded that the normal renal tissue destroyed the pressor substance, since otherwise its continued presence in the bloodstream should have kept the blood pressure elevated for a longer period. Dexter and Braun-Menendez (78) have demonstrated that the renal threshold for the excretion of renin is too high to account for the stabilizing rôle of the normal kidney. Subsequent studies by Grollman and collaborators (166) have led them to a modification of their original view and the adoption of the idea that the active principle derived from kidney tissue is an essential humoral agent, the absence of which, in the diseased kidney, results in hypertension. They now believe that the administration of this principle overcomes the deficiency and relieves the hypertension.

The demonstration that renal vein blood from an ischemic dog kidney (120) contains much less hypertensinase than renal vein blood from normal kidney has thrown light on the earlier experiments of Freeman (116) who found that normal dog's blood can reduce the blood pressure of dogs with experimental renal hypertension, whereas blood from hypertensive dogs has no such effect. The antipressor effect in this case may have been due to the presence of hypertensinase in normal blood and its diminution in the blood of hypertensive animals. These experiments require confirmation.

Page and collaborators (297, 311, 312, 313) and others who have used identical extracts (123, 281) have reported the reduction of arterial blood pressure of hypertensive animals and patients by the parenteral (intramuscular) administration of renal extract which contains a substance capable of inhibiting the pressor effect of angiotonin (59). Although Page and collaborators (313) have tried

to relate the antipressor effect of their kidney extract to the hypertensinase content, yet there is little hypertensinase in it, and Schales and collaborators (361) found that this extract gave the same result when its hypertensinase was first destroyed. Besides, there is no proof that hypertensinase *per se* is absorbed into the blood from an intramuscular injection.

It has been shown that for the antipressor effect of the renal extracts used for the treatment of hypertension, hypertensinase is not required. Harrison, Grollman and Williams (185) did not determine the existence of hypertensinase in their extract, but a lowering of blood pressure has been reported in the rat, dog and even man by the administration of renal extract both orally and intramuscularly. That such effects are caused by renin inhibitor, as postulated by Page and collaborators, has not been confirmed.

The possibility that all the effects on blood pressure by renal extracts, hitherto reported, are due to a non-specific pyrogenic effect (with or without actual elevation of temperature) of a foreign organic material that produces a local and general reaction, has already been suggested (46, 150). Indeed a reduction in the severity of the local reaction, as a result of purification, has resulted in a reduction of the antipressor effect. Schales, Stead and Warren (361) observed an antipressor effect which was not specific for renal extracts, and attributed all the effects they observed to the local and general reactions which occurred. They believe that this invalidates any conclusions about the specific effect of renal extracts. The same effect on blood pressure was obtained when the extract contained little or no hypertensinase, if the local and systemic reactions to the injections occurred. The extract that Grollman, Harrison and Williams used can hardly be the same as that of other investigators, because they found that the active substance was dialyzable and effective by mouth in both man and animals (174, 179, 180, 432). The experimental work on animals was all done on the rat, and has not been confirmed. No other investigators have reported a lowering of blood pressure in hypertensive individuals from the oral administration of renal extracts. Goldring and Chasis (153) failed to observe any favorable effect in four hypertensive patients treated with renal extract, given by mouth, up to the equivalent of 50 kgm. of original kidney substance daily.

Page and collaborators insist that they have observed fall of blood pressure from renal extract, without accompanying local or systemic reaction. However, it has been shown that all the other effects, except elevation of temperature, can be produced by the parenteral injection of foreign organic substances (46, 150, 153). The fact that there has been but little progress in the treatment of hypertension by renal extracts during the past 6 years is in itself an indication of the difficulties involved, and possibly of the inadequacy of the method.

Antirenin: Treatment of Hypertension by Means of Parenteral Injections of Renin. Because renin is a protein, it was to be expected that it would be antigenic and that parenteral injection of it would result in the development of an antibody in the blood plasma. What the properties of such an antiserum would be could not be anticipated; but it was hoped that it would prove to be antipressor, that is, antirenin.

In a series of studies, Wakerlin and collaborators (217, 414, 418, 419) reported that in the serum of rabbits, dogs and guinea pigs, but not the horse, injected with renin from various species, a substance or principle developed which neutralized the acute pressor effect of an intravenous injection of renin. They found that if the antiserum was first mixed with the renin and kept for about 18 hours at about 4°C, the mixture lost its pressor effect on the normal dog. Normal serum did not possess this property. The author and collaborators (223a) found subsequently that at room temperature the reaction between the antiserum and renin occurs practically instantaneously and that incubation for 18 hours is not required. Wakerlin and collaborators regard this principle as analogous to an antibody, antienzyme or antihormone, and suggested for it the name antirenin. This has now been produced in the rabbit, by the intramuscular injection of human, hog, cat and dog renin; in the guinea pig, by the intra-abdominal injection of hog, cat and dog renin, and in normal and renal hypertensive dogs, by the intramuscular injection of hog and rabbit renin. The injection of heat inactivated homologous or heterologous renin did not induce the formation of antirenin in either normal or hypertensive dogs (216, 415). Hog liver extract failed to produce antirenin in normal or hypertensive dogs.

Antirenin is a protein, but it is neither an enzyme nor a precipitin. Even when the antirenin titer is high, there is no precipitin in the blood of dogs or guinea pigs. In the rabbit, in which precipitin responses are more easily obtained, precipitins against other plasma proteins are formed in high titer when these animals are given repeated intravenous injections of renin. The serum of these animals still retains its ability to neutralize renin when the precipitin is removed by absorption *in vitro*. Antirenin is precipitated from serum by half saturation with ammonium sulphate. Antirenin is present in the pseudoglobulin fraction of the serum, where the circulating antibodies are ordinarily found. This fraction was found to contain practically all the antirenin of the original serum. It is active between a pH of 2.3 to 10.7 but is inactivated at a pH of 1.9 or less and 11.0 or more. At 0°C. and in the frozen state it remains stable for at least 3 months (223a). Antirenin does not neutralize the pressor effect of either hypertensin (217) or pituitrin (216). Like antihormones, (401) antirenin is not species specific in the neutralization of renin, *in vitro*, except for antirenin from rabbits induced by the injection of human renin. This specificity is also true for antihormones produced in rabbits by the injection of extracts from human tissue. Heating of the antiserum at 65°C for 10 minutes does not inactivate the antirenin, but the latter is destroyed by heating at 75°C. or higher (223a).

Wakerlin and collaborators were the first to report a fall of blood pressure to normal as a result of the repeated intramuscular injection of hog renin in hypertensive dogs. They also observed the prevention of hypertension in dogs that received similar injections before the renal arteries were constricted (415, 416). Both of these observations we have confirmed (223a). They found that the intraperitoneal injection of 40 ml. of dog serum containing a high titer of antirenin to hog renin had no specific effect on the blood pressure of one hypertensive dog.

No significant changes have been observed in the animals that have developed high titers of antirenin in their plasma. After a normal or hypertensive animal has developed a high titer of antirenin, large doses of homologous or heterologous renin may be injected intravenously into them without producing any change in the blood pressure. We have injected intravenously, in one dose, as much as 100 dog units of hog renin, without any elevation of the blood pressure, in a dog with 10 units of antirenin per cubic centimeter of its blood serum. The kidneys of rabbits and dogs which had a high titer of antirenin in the blood were found to contain normal amounts of renin (217).

Winternitz and collaborators (440) failed to confirm the production of antirenin, and Friedman and collaborators (133) did not observe a fall of blood pressure in animals that received parenteral injections of renin. But the former injected the renin intravenously and the latter probably injected an insufficient amount. The author and collaborators also failed to detect the production of antirenin in dogs by the intravenous injection of 25 units of hog renin a day, for 12 weeks, but, by subcutaneous and intramuscular injection of the same amount of hog renin, we did succeed in producing antirenin of high titer in the plasma of normal and hypertensive dogs. We also observed a fall of blood pressure in hypertensive dogs and prevented the development of hypertension in dogs by repeated subcutaneous injections of hog renin before the renal arteries were constricted. The optimum dose was 25 units of renin daily.

Wakerlin and collaborators were the first to produce antirenin, yet they now consider that they have good evidence that the antirenin is not responsible for the prevention of hypertension or the lowering of the blood pressure in dogs with renal hypertension. The evidence upon which they repudiate the significance of the antirenin is not conclusive. One reason given is that highly purified hog renin was as effective as partially purified hog renin in stimulating the formation of antirenin, but that it was much less effective as an antihypertensive agent. It is highly probable, however, that this difference was due only to the smaller amount of renin injected. Another reason was that in an occasional dog injected with hog renin inactivated by heat, a fall in blood pressure was noted. Not too much significance should be attached to this, because in some untreated dogs the blood pressure also returns to a lower level. They also found no correlation between the antirenin titer and the antihypertensive effect. This is contrary to the results of our own studies. Because they did not give an exact determination of the titer of antirenin attained by their animals, it is difficult to make a direct comparison between their results and ours. The methods of assay of antirenin are not comparable. Their antirenin titer is based on the amount of renin extracted per gram of renal cortex, which is obviously subject to considerable variation. We titrate the exact amount of antirenin by determining the number of units of renin inactivated by the serum. Our dog unit of antirenin is the minimum quantity which neutralizes the acute pressor effect of 1 dog unit of renin.

The possible application of the results obtained in animals to the treatment of human hypertension is beset with difficulties. Because homologous renin

does not produce antirenin, the injection of human renin into hypertensive patients would be ineffective. Whether heterologous animal renins which do not produce a pressor effect in man would induce the formation of antirenin is still unknown. Much more work on this subject is required. It has not been given adequate attention by other investigators.

Non-specific Pressor and Antipressor Effects. In any consideration of the renal humoral mechanism, mention must be made of the studies by Bing and collaborators (19, 20, 21) and others (205) which, although they may have no direct bearing on this mechanism, yet indicate a possible way whereby a pressor substance of renal origin may be formed, or released, under conditions which involve a disturbance of intrarenal hemodynamics. They demonstrated (20) by perfusion experiments that an ischemic kidney, but not a normal one, is capable, by decarboxylation, of converting dihydroxyphenylalanine (dopa), a substance which possesses no pressor properties, into hydroxytyramine, which is a powerful pressor amine. They showed that the amount of hydroxytyramine which is formed in the kidney from l-dopa under conditions of oxygen lack was proportional to the reduction of blood flow through the perfused kidney (19). They also demonstrated (19) the transformation of dopa into hydroxytyramine in kidney extract under conditions of low oxygen tension. But even without the addition of l-dopa the development of a pressor substance has been reported merely as a result of anaerobic conditions (407). Liver, spleen, lung and heart treated the same way did not develop pressor substances. Liver and intestine also contain l-dopa decarboxylase, yet these organs were unable, even when their circulation was reduced, to produce hydroxytyramine from l-dopa which was added to the blood (19).

Bing was also able to show that partly or completely ischemic cat's kidneys, *in vivo*, converted l-dopa, injected intravenously, into hydroxytyramine. The observation of Oster and Soskin (291) that the intravenous injection of l-dopa into cats with experimental renal hypertension resulted in a great rise of blood pressure, while no rise occurred in normal cats, is of interest in this connection. But Bing and collaborators did not conclude that hydroxytyramine is the cause of, or in any way directly involved in, the pathogenesis of experimental renal or human hypertension. The same conclusion has been reached about tyramine (100, 347, 404). Bing and collaborators have demonstrated that the two pressor substances, hydroxytyramine and hypertensin, are destroyed by different fractions of renal extract and by different mechanisms. It has also been shown that renin does not effect the decarboxylation of l-dopa to convert it to hydroxytyramine (67). It has been demonstrated that the reduction of the blood flow to the kidney results in a profound alteration of the subsequent chemical events. The kidney has the capability of converting the amino acid, which is itself without pressor property, into a powerful pressor amine which accumulates under anaerobic conditions and results in elevation of blood pressure. Pressor amines are rapidly destroyed, however, by oxidative enzymes when the circulation is normal and aerobic conditions prevail. Their work has provided a basis for the hypothesis that the hypertension which results from renal ischemia may be due to

diminished deamination of certain amino acids in the kidney and that it may possibly play an adjuvant or even an important primary part in hypertension in some types of animals.

The work of Bing and collaborators has been the basis of the attempts (127, 290, 386) at the treatment of experimental renal hypertension in rats by means of the quinones, and, in hypertensive dogs, by a diketone, 1-4-cyclohexandione (288). A lowering of blood pressure was reported, but we have not been able to confirm any of these results by the use of these materials in hypertensive dogs.

Pressor amines certainly exist and produce their effect when injected intravenously into animals. Their deamination requires oxygen. The question is whether some disturbance in the kidney may so reduce the available oxygen that deamination of natural amines does not occur and that the entrance of such amines into the blood stream results in hypertension. There is, as yet, no available evidence from the studies on man that this mechanism plays a part in human hypertension. However, the studies on the treatment of human hypertension with amine oxidases (365) has not led to an elucidation of this problem. Schroeder and collaborators (362, 366, 367) found that the parenteral injection of tyrosinase, a phenol oxidase, reduced the blood pressure in rats and dogs with experimental renal hypertension, and in man, with essential hypertension. But Prinzmetal and collaborators (332) showed that the effect obtained by Schroeder was not due to the enzymatic acitvity of the tyrosinase, because the extract containing it was also effective when the tyrosinase activity was first destroyed.

The observation that both hypertensin and pepsitensin are inactivated by an amino-peptidase obtained from yeast has led Croxatto and Croxatto (69) to conclude that the hypertensinase activity of renal extract may also be attributed to the enzyme amino-peptidase contained in renal tissue. They have also shown (66) that the vasoconstrictor effect of hypertensin tested on the perfused toad is destroyed enzymatically by the action of amine oxidase and tyrosinase. In our own laboratory, in collaboration with Gollan and Richardson, much work has been done on enzymes, from many plants that are capable of inactivating hypertensin *in vitro*. It is difficult to make preparations that are not toxic, and although antipressor effects have been obtained in hypertensive dogs, yet it is difficult to estimate how much of the effect was due specifically to the "plant hypertensinase".

The problem whether there is a renal mechanism of hypertension which is dependent upon a metabolic fault in the kidney, interfering with the utilization of melanin-like substances and their phenolic precursors, is not yet settled. Even the possibility that such a mechanism exists in one animal, the rat, for example, and not in other animals, including man, has not been determined with certainty.

The effect of diet on experimental renal hypertension has not been exhaustively investigated. It has been asserted that a high protein (800 g. meat daily) diet, or 50 g. of urea daily, increases the elevated blood pressure in dogs with experimental renal hypertension (259). A diet that produces considerable gain of weight has a similar effect (43).

The significance of the experiments of Calder (40, 41) on the production of hypertension in the rat by a diet deficient in the heat stable fraction of the vitamin B complex is of interest. His belief is that the hypertension is of metabolic origin and due to diminished oxidative processes in the kidney. This work requires confirmation.

The claim for the effectiveness of ascorbic acid in reducing the blood pressure of hypertension in man (75) has not been confirmed (276). This has not been tested on hypertensive animals. It has been reported that vitamin E has no effect on the blood pressure of dogs with experimental renal hypertension (276). This has not been tested on hypertension in man. The effect of a special rice diet (231, 232) in lowering the blood pressure of human hypertensives has not been substantiated, and the mechanism of its action is not clear. This has not been tested on hypertensive animals. The assertion that large doses of vitamin A lower the blood pressure of hypertensive human beings (159), rats (169) and dogs (420, 421) has not been substantiated by the same investigators (159, 168, 417), who finally decided that it was not the vitamin A that produced the effect (164, 168, 275, 276, 417) because fish body and liver oil still contained the blood pressure reducing substance when the vitamin A was destroyed. In fact, Grollman and collaborators (417) found that a highly purified vitamin A concentrate did not have the hypotensive effect and that the effect was enhanced by the oxidation of the fatty acid in effective marine oil. The nature and mechanism of the action of this substance in marine oils is still unknown.

In his most recent publication on this subject, Grollman (165) asserts that a number of refined oils derived from marine fishes, when administered orally, reduce blood pressure in hypertensive rats, dogs and man. Tung oil, but no other oils of vegetable or animal origin, was also effective in hypertensive animals. Oxidation, and saponification prior to oxidation, enhanced the activity of some of the oils. The effect was independent of their original vitamin A content and was retained after the destruction of this vitamin. The active principle is soluble in water and dialyzable, as in the case of the orally effective renal extracts (185, 187) with which he is inclined to consider this identical. He states, however, that it is still impractical to treat patients because there is no readily available source of large amounts of this principle.

It has been reported (282) that vitamin A may affect urea and inulin clearances, but these authors did not control their experiments by giving the same product with the vitamin A inactivated. Vitamin D_2 given by mouth had no effect on the blood pressure of the rat (38). Another sterol that has no effect on experimental renal hypertension in the dog is testosterone (412).

On the basis of experiments on hypertensive rats, a disequilibrium in sodium balance was noted by Grollman and collaborators (167, 170, 171) which led them to test the effect of restriction of sodium intake on hypertension in the rat and in man. They noted a lowering of blood pressure in both. The significance of these observations remains to be elucidated.

The most recent contribution by Shorr, Zweifach and Furchgott (380) on the humoral mechanism of shock is of great interest, but the results must await confirmation and elucidation before they can be evaluated. The identity of the

vaso-excitor (VEM) and vaso-depressor (VDM) substances which they have described, with the known constituents of the humoral mechanism of renal hypertension, cannot be considered as established on the basis of the evidence submitted. The possible part these substances may play in the maintenance of blood pressure at normal and elevated levels cannot yet be determined from the results that have been published.

A Summary of the Similarities Between Human Essential and Experimental Renal Hypertension. Although Goldring and Chasis in their excellent book (153) conclude that the weight of the evidence is against the identity of the mechanism in human and experimental renal hypertension, yet it has been shown by many investigators that the latter does faithfully reproduce human essential hypertension in many respects. In both, the increased tension is the result of a generalized increase of peripheral vascular resistance of functional (vasospastic) origin. In experimental renal hypertension, as in human hypertension, there may be no significant disturbance of renal excretory function (the benign phase) or there may be pronounced renal excretory functional disturbance, with uremia (the malignant phase), depending entirely upon the degree of constriction of the main renal arteries (144). As in human hypertension, so also in experimental renal hypertension, cardiac action is increased (258), but cardiac rate, output (204), volume (15), viscosity (297) and peripheral flow of the blood (221), and venous pressure remain unaltered. Pulmonic arterial pressure is not altered in either man or animals, when the hypertension is uncomplicated by left ventricular failure (227, 391), as indicated by a normal right heart. In both man and animals hypertension associated with unilateral renal disease may be cured by excision of the diseased kidney, provided the other is normal (252). Bilateral nephrectomy does not result in a rise in blood pressure in either man or animal (144). With few exceptions, the response to medicinal therapeutic measures of great variety is the same in both. Sympathectomy, partial or extensive, may result in at least a temporary fall of blood pressure in human hypertension, yet it does not do so by affecting the primary cause of the hypertension; but there is little or no effect in animals that do not stand erect. Whether this difference is significant cannot be stated with certainty at present. The fact that, after sympathectomy, the blood pressure returns to the original high level in a high percentage of the hypertensive patients also favors the idea of a renal humoral mechanism in which the effective vasoconstrictor substance is presumed to act directly on the musculature of the peripheral arterioles, and not by way of the vasomotor nerves. This is in keeping with the conclusions of Prinzmetal and Wilson (337) and of Pickering (320). The frequent fall of blood pressure in the late stage of pregnancy in animals with experimental renal hypertension (58, 77, 86, 87, 111, 292) remains as unexplained as does a similar fall which has been observed by many obstetricians in some hypertensive pregnant women. The observation of polydypsia and polyuria, in rats with experimental renal hypertension (289), has not been emphasized in human hypertension and has not been observed in hypertensive dogs, but the diuretic effect of renin injected intravenously into animals has been mentioned (272). This should be investigated further.

Renal blood flow is reduced in most cases of human hypertension (135, 154 384) and in experimental renal hypertension (306) in animals. The indirect studies (135, 154, 384) of blood flow through the kidney in man do not demonstrate clearly the primary effect of the sclerosis of the afferent arterioles, because the vasospasm of efferent arterioles which results in high glomerular filtration fraction tends to mask it. Interestingly enough, although the interference with afferent flow is definite, in the animals, and obviously brought about by constriction of the main renal artery, yet the same indirect signs of efferent vasospasm and high glomerular filtration fraction occur in the hypertensive animals. In both, the latter effects may be due to the humoral mechanism of renal origin. In both man and animals the hypertension may or may not be accompanied by disturbance of renal excretory function. An increase in the concentration of guanidine in the blood of animals (49, 51) and man (260, 261), in the malignant phase of hypertension, has been demonstrated; but it has little or no significance with relation to the hypertension, because it occurs also in bilaterally nephrectomized animals that have the azotemia but not the hypertension (80). The presence of renin has been demonstrated in the renal venous blood of ischemic kidneys of man and animals. Although renin has been demonstrated in the systemic blood of recently hypertensive dogs (215), and man with hypertension due to acute glomerulonephritis, yet the failure to demonstrate it in the systemic blood of dogs with long standing renal hypertension and patients with benign essential hypertension may be only because the amount of blood used for the tests has been inadequate and the methods have lacked sensitivity. Whether the humoral mechanism is effective only in the relatively acute stages of hypertension, or whether in the later stages, as has been suggested (39, 285), there is a greatly increased sensitivity to hypertensin, remains to be determined. These matters deserve much more study. In the benign phase of hypertension, in both man and animals, cardiac hypertrophy develops, affecting mainly the left ventricle, and medial hypertrophy of the arterial vessels also occurs in both (252). In the malignant phase, in both, there are the identical typical vascular lesions, arteriolar necrosis, fibrinoid degeneration and necrotizing arteriolitis (147). As in human, so also in experimental renal hypertension, the level of blood pressure tends to go to a higher level in hypertensive animals which gain weight (43), but a high protein diet or the ingestion of a large amount of urea has not been found to increase the blood pressure of hypertensive dogs (183). Although the elevated systolic blood pressure of hyperthyroidism is relieved by thyroidectomy, yet it is doubtful that any of the known endocrine organs plays a primary part in either essential hypertension associated with vascular disease, in man, or in experimental renal hypertension in animals. There are definite indications that the adrenal cortical hormones may play a secondary part in the development and maintenance of experimental renal and human essential hypertension (85, 125, 128, 318, 319, 342).

It would be remarkable indeed if, despite all these close similarities between human essential hypertension, associated with vascular disease, and a type of experimental hypertension that is obviously of renal origin, the former would

not prove to be of renal origin. In the event that the renal origin of this form of human hypertension should become established, it would still be necessary to determine the cause of the arterial and arteriolar sclerosis which, when it affects the kidneys to a sufficient degree, initiates the humoral mechanism of the hypertension. It is still a fascinating subject for investigation and much remains to be done.*

* Recently there was published a translation, in English, by Dexter, of the book on Renal Hypertension by Braun-Menendez and collaborators (37a). In the prologue, Houssay has written, "In the clinic, the existence of hypertension from renal ischemia, similar to experimental renal hypertension has not only been demonstrated, but the hypertension has been cured by removal of the diseased kidney in some cases of unilateral renal disease.

REFERENCES

(1) Abeshouse, B. S. Surgery 9: 942, 1941; 10: 147, 1941.
(2) Abramson, D. I. and B. Friedman. Am. J. Physiol. 123: 1, 1938.
(3) Allbutt, T. C. Arteriosclerosis: A summary view. Macmillan & Co., N. Y., 1925.
(4) Allen, E. V. Proc. Staff Meet. Mayo Clinic 17: 519, 1942.
(5) Allen, F. M. J. Urol. 49: 512, 1943.
(6) Alonso, O., R. Croxatto and H. Croxatto. Proc. Soc. Exper. Biol. and Med. 52: 61, 1943.
(7) Alpert, L. K., A. S. Alving and K. S. Grimson. Proc. Soc. Exper. Biol. and Med. 37: 1, 1937.
(8) Anderson, E., E. W. Page, C. H. Li and E. Ogden. Am. J. Physiol. 141: 393, 1944.
(9) Apfelbach, C. W. and C. R. Jensen. J. Clin. Investigation 10: 162, 1931.
(10) Bobrös, K. J. A. M. A. 119: 914, 1942.
(11) Baggenstoss, A. H. and N. W. Barber. Arch. Path. 32: 966, 1941.
(12) Barthels, E. C., J. L. Poppen and R. L. Richards. Ann. Int. Med. 17: 807, 1942.
(13) Blau, J. W. Federation Proc. 1: 6, 1942.
(14) Blau, J. W. Am. J. Physiol. 136: 731, 1942.
(15) Beckwith, J. R. and A. Chanutin. Proc. Soc. Exper. Biol. and Med. 46: 66, 1941.
(16) Bell, E. T. and B. J. Clawson. Arch. Path. 5: 939, 1928.
(17) Bell, E. T. and A. H. Pedersen. Ann. Int. Med. 4: 227, 1930.
(18) Besson, J. H. Northwest Med. 43: 78, 1944.
(19) Bing, R. J. Am. J. Physiol. 132: 407, 1941.
(20) Bing, R. J. and M. B. Zucker. Am. J. Physiol. 133: 214, 1941.
(21) Bing, R. J. and M. B. Zucker. J. Exper. Med. 74: 235, 1941.
(22) Bingel, A. and E. Strauss. Deutsch. Arch. f. klin. Med. 96: 476, 1909.
(23) Blackman, S. S., Jr. Bull. Johns Hopkins Hosp. 65: 353, 1939.
(24) Blalock, A. Physiol. Rev. 20: 159, 1940.
(25) Blalock, A. and S. E. Levy. Ann. Surg. 106: 826, 1937.
(26) Bouckaert, J. J., K. S. Grimson and C. Heymans. Am. J. Physiol. 96: 44P, 1939.
(27) Bowen, B. D. and N. Kutzman. Ann. Int. Med. 17: 427, 1942.
(28) Boylston, G. A., E. G. McEwen and A. C. Ivy. Proc. Soc. Exper. Biol. and Med. 39: 559, 1938.
(29) Braasch, W. F., W. Walters and H. J. Hammer. Proc. Staff Meet. Mayo Clin. 15: 477, 1940.
(30) Braasch, W. F. and W. W. Wood. J. Urol. 48: 343, 1942.
(31) Braun-Menendez, E. Ann. Rev. Physiol. 6: 265, 1944.
(32) Braun-Menendez, E. and J. C. Fasciolo. Compt. rend. Soc. Biol. 133: 324, 1940.
(33) Braun-Menendez, E., J. C. Fasciolo, L. F. Leloir and J. M. Muñoz. Compt. rend. Soc. de Biol. 133: 728, 1940.

(34) Braun-Menendez, E., J. C. Fasciolo, L. F. Leloir and J. M. Muñoz. Rev. Soc. argent. de biol. **15**: 420, 1939.
(35) Braun-Menendez, E., J. C. Fasciolo, L. F. Leloir and J. M. Muñoz. Compt. rend. Soc. de Biol. **133**: 731, 1940.
(36) Braun-Menendez, E., J. C. Fasciolo, L. F. Leloir and J. M. Muñoz. Am. J. Physiol. **98**: 283, 1940.
(37) Braun-Menendez, E., J. C. Fasciolo, L. F. Leloir, J. M. Muñoz and A. C. Taquini. Rev. Soc. argent. de biol. **19**: 304, 1943.
(37a) Braun-Menendez, E., Fasciolo, J. C., Leloir, L. F., Muñoz, I. M., Taquini, A. C. (Translated by Dexter, L.) Renal Hypertension. Charles C. Thomas, Springfield, Ill., U. S. A., 1946.
(38) Briskin, H. L., F. R. Stokes, C. I. Reed and R. G. Mrazek. Am. J. Physiol. **138**: 385, 1943.
(39) Brown, G. M. and B. G. Maegraith. J. Physiol. **99**: 304, 1940.
(40) Calder, R. M. J. Exper. Med. **76**: 1, 1942.
(41) Calder, R. M. J. Exper. Med. **79**: 215, 1944.
(42) Cash, J. R. Bull. Johns Hopkins Hosp. **35**: 168, 1924.
(43) Cash, J. R. and J. E. Wood. South. Med. J. **31**: 270, 1938.
(44) Castleman, B. and R. H. Smithwick. J. A. M. A. **121**: 1256, 1943.
(45) Cerqua, S. and A. Samaan. Clin. Sc. **4**: 113, 1939.
(46) Chasis, H., W. Goldring and H. W. Smith. J. Clin. Investigation **21**: 369, 1942.
(47) Chasis, H. and J. Redish. J. Clin. Investigation **20**: 655, 1941.
(48) Chasis, H. and J. Redish. Arch. Int. Med. **70**: 738, 1942.
(49) Child, C. G. Arch. Path. **25**: 768, 1938.
(50) Child, C. G. and F. Glenn. Arch. Surg. **36**: 376, 1938.
(51) Child, C. G. J. Clin. Investigation **17**: 301, 1938.
(52) Collings, W. D., J. W. Remington, H. W. Hays and V. A. Drill. Proc. Soc. Exper. Biol. and Med. **44**: 87, 1940.
(53) Collins, D. A. Am. J. Physiol. **116**: 616, 1936.
(54) Collins, D. A. and A. S. Hamilton: Am. J. Physiol. **129**: 336, 1940.
(55) Collins, D. A. and A. S. Hamilton. Am. J. Physiol. **130**: 784, 1940.
(56) Collins, D. A. and A. S. Hamilton. Am. J. Physiol. **140**: 499, 1943.
(57) Collins, D. A. and E. H. Wood. Am. J. Physiol. **123**: 224, 1938.
(58) Corbit, J. D., Jr. Am. J. Med. Sc. **201**: 876, 1941.
(59) Corcoran, A. C., K. G. Kohlstaedt and I. H. Page. Am. J. Physiol. **133**: 248, 1941.
(60) Corcoran, A. C. and I. H. Page. Am. J. Physiol. **123**: 43, 1938.
(61) Corcoran, A. C. and I. H. Page. Am. J. Physiol. **130**: 335, 1940.
(62) Corcoran, A. C. and I. H. Page. Am. J. Physiol. **133**: 249, 1941.
(63) Corcoran, A. C. and I. H. Page. J. A. M. A. **116**: 690, 1941.
(64) Corrigan, F. P. and I. Pines. Surgery **14**: 88, 1943.
(65) Cox, A. J. and W. Dock. J. Exper. Med. **74**: 167, 1941.
(66) Croxatto, H. and R. Croxatto. Proc. Soc. Exper. Biol. and Med. **48**: 392, 1941.
(67) Croxatto, R. and H. Croxatto. Rev. de med. y. aliment. **5**: 66, 1942.
(68) Croxatto, H. and R. Croxatto. Science **95**: 101, 1942.
(69) Croxatto, R. and H. Croxatto. Science **96**: 519, 1943.
(70) Croxatto, H., R. Croxatto, H. Manriquez and B. Valenzuela. Rev. de med. y. aliment. **5**: 137, 1942.
(71) Croxatto, R., H. Croxatto and L. Marty. Proc. Soc. Exper. Biol. and Med. **52**: 64, 1943.
(72) Croxatto, R., H. Croxatto and G. Rodriguez. Rev. Soc. argent. de biol. **21**: 172, 1945.
(73) Cruz Coke, E. Conference on Experimental Hypertension, The N. Y. Acad. of Sciences, Section of Biology, Feb. 9–10, 1945.
(74) Dalton, J. W. and F. R. Nuzum. Arch. Int. Med. **70**: 948, 1942.

(75) Davis, N. S. and E. F. Poser. J. A. M. A. **122**: 59, 1943.
(76) Davson, J. J. Path. and Bact. **53**: 207, 1941.
(77) Dawson, J. R., R. D. Cressman and A. Blalock. Am. J. Path. **17**: 31, 1941.
(78) Dell'Oro, R. Rev. de med. y. aliment. **5**: 126, 1942.
(79) Dell'Oro, R. Rev. Soc. argent. de. biol. **18**: 13, 1942.
(80) Dell'Oro, R. and E. Braun-Menendez. Rev. Do la Soc. Argentina de Biol. **18**: 65, 1942.
(81) Dexter, L. Ann. Int. Med. **17**: 147, 1942.
(82) Dexter, L. and F. W. Haynes. Proc. Soc. Exper. Biol. and Med. **55**: 288, 1944.
(83) Dexter, L., H. A. Frank, F. W. Haynes and M. D. Altschule. J. Clin. Investigation **22**: 847, 1943.
(84) Dexter, L., F. W. Haynes and W. C. Bridges. J. Clin. Investigation **24**: 62, 1945.
(85) Diaz, J. T. and S. E. Levy. Am. J. Physiol. **125**: 586, 1939.
(86) Dill, L. V. and C. C. Erickson. Arch. Path. **31**: 68, 1941.
(87) Dill, L. V., C. E. Isenhour, J. F. Cadden and A. Kuder. Surg., Gynec. and Obst. **72**: 38, 1941.
(88) Dock, W. Am. J. Physiol. **130**: 1, 1940.
(89) Dock, W. J. Clin. Investigation **19**: 769, 1940.
(90) Dock, W. Kansas City Med. J. (May–June) 8, 1942.
(91) Drill, V. A. Proc. Soc. Exper. Biol. and Med. **49**: 557, 1942.
(92) Drury, D. R. J. Exper. Med. **68**: 693, 1938.
(93) Duff, G. L., J. D. Hamilton and D. Magner. Proc. Soc. Exper. Biol. and Med. **41**: 295, 1939.
(94) Dunihue, F. W. Arch. Path. **32**: 211, 1941.
(95) Dunihue, F. W. and B. H. Candon. Arch. Path. **29**: 777, 1940.
(96) Edman, P., U. S. Euler, E. Jorpes and O. T. Sjostrand. J. Physiol. **101**: 284, 1942.
(97) Edman, P. Arkiv. för Kemi, Mineralogioch Geologi, **22**: 1, 1945.
(98) Eichelberger, L., L. Leiter and E. M. K. Geiling. Proc. Soc. Exper. Biol. and Med. **44**: 356, 1940.
(99) Levaditi, C. and R. Schoen. Compt. rend. Soc. de biol. **122**: 616, 1936.
(100) Elsom, K. A. and P. M. Glenn. Ann. Int. Med. **12**: 838, 1938.
(101) Emmel, V. M. Anat. Rec. **78**: 361, 1940.
(102) Enger, R., F. Linder and H. Sarre. Ztschr. f.d. ges. exper. Med. **104**: 1, 1938.
(103) Enger, R., F. Linder and H. Sarre. Ztschr. f.d. ges. exper. Med. **104**: 10, 1938.
(104) v. Euler, U. S. and O. T. Sjöstrand. Acta physiol. Scand. **5**: 183, 1943.
(105) Fasciolo, J. C., B. A. Houssay and A. C. Taquini. J. Physiol. **94**: 281, 1938.
(106) Fasciolo, J. C. Compt. rend. Soc. de biol. **128**: 1120, 1938.
(107) Fasciolo, J. C., L. F. Leloir, J. M. Muñoz and E. Braun-Menendez. Rev. Soc. argent. de biol. **16**: 643, 1940.
(108) Findley, T., J. C. Edwards, E. Clinton and H. L. White. Arch. Int. Med. **70**: 935, 1942.
(109) Fishberg, A. M. Arch. Int. Med. **35**: 650, 1925.
(110) Fishberg, A. M. J. A. M. A. **119**: 551, 1942.
(111) Foà, P. P., N. L. Foà and M. M. Peet. Am. J. M. Sc. **204**: 350, 1942.
(112) Foà, P. P., W. W. Woods, M. M. Peet and N. L. Foà. Arch. Int. Med. **71**: 357, 1943.
(113) Foglia, V. G. and J. L. Moglia. Rev. Soc. argent. de biol. **16**: 520, 1940.
(114) Frankel, D. B. and G. E. Wakerlin. Am. J. Physiol. **138**: 465, 1942.
(115) Freedman, A. Am. Heart J. **20**: 304, 1940.
(116) Freeman, G. Proc. Soc. Exper. Biol. and Med. **45**: 185, 1940.
(117) Freeman, N. E. and I. H. Page. Surgery **2**: 487, 1937.
(118) Friedberg, L. Am. J. Physiol. **133**: 282, 1941.
(119) Freeman, N. E. and W. A. Jeffers. Am. J. Physiol. **126**: 493, 1939.
(120) Friedberg, L., M. Landowne and S. Rodbard. Am. J. Physiol. **129**: 358, 1940.
(121) Friedman, B., D. I. Abramson and W. Marx. Am. J. Physiol. **124**: 285, 1938.

(122) Friedman, B., J. Jarman and P. Klemperer. Am. J. M. Sc. **202**: 20, 1941.
(123) Friedman, B., J. Jarman and J. Marrus. J. Mt. Sinai Hosp. **8**: 534, 1941.
(124) Friedman, B., L. Moschkowitz and J. Marrus. J. Urol. **48**: 5, 1942.
(125) Friedman, B., B. S. Oppenheimer, E. Sompkin, E. T. Oppenheimer and B. Blumenthal. J. Clin. Investigation **18**: 477, 1939.
(126) Friedman, B. and M. Prinzmetal, Ann. Int. Med. **12**: 1617, 1939.
(127) Friedman, B., S. Soloway, J. Marrus and B. S. Oppenheimer. Proc. Soc. Exper. Biol. and Med. **51**: 195, 1942.
(128) Friedman, B., E. Somkin and E. T. Oppenheimer. Am. J. Physiol. **128**: 481, 1939.
(129) Friedman, M. Proc. Soc. Exper. Biol. and Med. **47**: 348, 1941.
(130) Friedman, M. and A. Kaplan. J. Exper. Med. **75**: 127, 1942.
(131) Friedman, M. and A. Kaplan. J. Exper. Med. **77**: 65, 1943.
(132) Friedman, M., A. Kaplan and E. Williams. Proc. Soc. Exper. Biol. and Med. **50**: 199, 1942.
(133) Friedman, M., H. E. Kruger and A. Kaplan. Am. J. Physiol. **137**: 570, 1942.
(134) Friedman, M., W. Marx and E. Lindner. Proc. Soc. Exper. Biol. and Med. **54**: 221, 1943.
(135) Friedman, M., A. Selzer and H. Rosenblum. J. A. M. A. **117**: 92, 1941.
(136) Friedman, M., A. Selzer and J. J. Sampson. Am. J. Physiol. **131**: 799, 1940.
(137) Gabriele, D. J. Am. J. Med. Sci. **204**: 227, 1942.
(138) Gaudino, N. M. Rev. Soc. argent. de biol. **20**: 470, 1944.
(139) Geer, W. A. and L. R. Dragstedt. Ann. Surg. **108**: 263, 1938.
(140) Gerbi, C., B. B. Rubinstein and H. Goldblatt. J. Exper. Med. **71**: 71, 1940.
(141) Gibson, T. E. Calif. and West. Med. **56**: 66, 1942.
(142) Glenn, F., C. G. Child and G. J. Heuer. Ann. Surg. **106**: 848, 1937.
(143) Glenn, F. and E. P. Lasher. Am. J. Physiol. **124**: 106, 1938.
(144) Goldblatt, H. The Harvey Lectures, 1937–1938, pp. 237–275.
(145) Goldblatt, H. J. Exper. Med. **67**: 809, 1938.
(146) Goldblatt, H. J. Clin. Path. **10**: 40, 1940.
(147) Goldblatt, H. Am. Assn. Adv. Sci., no. 13, 266, 1940.
(148) Goldblatt, H. Nelson Loose Leaf, 1941.
(149) Goldblatt, H. Medical physics, p. 622, 1944. The Year Book Publishers.
(150) Goldblatt, H., J. R. Kahn and H. A. Lewis. J. A. M. A. **119**: 1192, 1942.
(151) Goldblatt, H., Y. J. Katz, H. A. Lewis and E. Richardson. J. Exper. Med. **77**: 309, 1943.
(151a) Goldblatt, H., J. R. Kahn, H. A. Lewis, E. Richardson, A. Guevara-Rojas and F. Gollan. J. A. M. A. **122**: 135, 1943.
(152) Goldblatt, H., J. Lynch, R. F. Hanzal and W. W. Summerville. J. Exper. Med. **59**: 347, 1934.
(153) Goldring, W. and H. Chasis. Hypertension and hypertensive disease. The Commonwealth Fund, 1944.
(154) Goldring, W., H. Chasis, H. A. Ranges and H. W. Smith. J. Clin. Investigation **17**: 505, 1938.
(155) Goldring, W., H. Chasis, H. A. Ranges and H. W. Smith. J. Clin. Investigation **20**: 637, 1941.
(156) Goormaghtigh, N. Am. J. Path. **16**: 409, 1940.
(157) Goormaghtigh, N. Rev. belge Sci. méd. **12**: 85, 1940.
(158) Goormaghtigh, N. and K. Grimson. Proc. Soc. Exper. Biol. and Med. **42**: 227, 1939.
(158a) Goormaghtigh, N. La fonction endocrine des artérioles rénales. Libraire R. Fonteyn, Louvain, 1944.
(159) Govea-Peña, J. and M. Villanerde. A. Rev. Cubana Cardiol. **2**: 322, 1940.
(160) Graef, I. and H. W. Smith. J. Clin. Investigation **19**: 770, 1940.
(161) Greenwood, W. F., R. Nassim and N. B. Taylor. Can. M. A. J. **41**: 443, 1939.

(162) Gregory, R., P. L. Ewing, W. C. Levin and G. T. Ross. Arch. Int. Med. **76:** 11, 1945.
(163) Grimson, K. S. Am. J. Physiol. **95:** 45P, 1939.
(164) Grollman, A. Fed. Proc. **3:** 15, 1944.
(165) Grollman, A. J. Pharmacol. and Exper. Therap. **84:** 128, 1945.
(166) Grollman, A. The N. Y. Acad. of Science, Section of Biology, Feb. 9–10, 1945
(167) Grollman, A. Texas State J. Med. **41:** 304, 1945.
(168) Grollman, A. and T. R. Harrison. Fed. Proc. **2:** 81, 1943.
(169) Grollman, A. and T. R. Harrison. Proc. Soc. Exper. Biol. and Med. **52:** 162, 1943.
(170) Grollman, A. and T. R. Harrison. Proc. Soc. Exper. Biol. and Med. **60:** 52, 1945.
(171) Grollman, A., T. R. Harrison, M. F. Mason, J. Baxter, J. Crampton and F. Reischman. J. A. M. A. **129:** 533, 1945.
(172) Grollman, A., T. R. Harrison and J. R. Williams, Jr. Am. Assn. Adv. Sci., no. 13: 274, 1940.
(173) Grollman, A., T. R. Harrison and J. R. Williams, Jr. J. Pharmacol. and Exper. Therap. **69:** 76, 1940.
(174) Grollman, A., T. R. Harrison and J. R. Williams, Jr. J. Pharmacol. and Exper. Therap. **69:** 149, 1940.
(175) Grollman, A., T. R. Harrison and J. R. Williams, Jr. Fed. Proc. **1:** 34, 1942.
(176) Grollman, A., T. R. Harrison and J. R. Williams, Jr. Am. J. Physiol. **139:** 293, 1943.
(177) Grollman, A. and C. Rule. Am. J. Physiol. **138:** 587, 1942.
(178) Grollman, A. and J. R. Williams, Jr. Am. J. Med. Sc. **204:** 73, 1942.
(179) Grollman, A., J. R. Williams, Jr. and T. R. Harrison. J. A. M. A. **115:** 1169, 1940.
(180) Grollman, A., J. R. Williams, Jr. and T. R. Harrison. J. Biol. Chem. **134:** 115, 1940.
(181) Grossman, E. B. Proc. Soc. Exper. Biol. and Med. **39:** 40, 1938.
(182) Grossman, E. B. and J. R. Williams, Jr. Arch. Int. Med. **62:** 799, 1938.
(183) Guerrant, J. L., J. K. Scott, and J. E. Wood, Jr. Am. Heart J. **26:** 232, 1943.
(184) Hamilton, A. S. and D. A. Collins. Am. J. Physiol. **136:** 275, 1942.
(185) Harrison, T. R., A. Grollman and J. R. Williams, Jr. Am. J. Physiol. **128:** 716, 1939.
(186) Harrison, T. R., A. Grollman and J. R. Williams, Jr. Trans. Assn. Am. Phys. **57:** 187, 1942.
(187) Hartwich, A. and G. Hessel. Zentralbl. inn. Med. **18a:** 612, 1932.
(188) Hartwich, A. and G. Hessel. Zentralbl. inn. Med. **18a:** 626, 1932.
(189) Haynes, F. W. and L. Dexter. Fed. Proc. **2:** 20, 1943.
(190) Haynes, F. W. and L. Dexter. J. Clin. Investigation **24:** 75, 1945.
(191) Helmer, O. M. and I. H. Page. J. Biol. Chem. **127:** 757, 1939.
(192) Helmer, O. M. and I. H. Page. Proc. Soc. Exper. Biol. and Med. **49:** 389, 1942.
(193) Herrick, J. F., A. C. Corcoran and H. E. Essex. Am. J. Physiol. **133:** 324, 1941.
(193a) Herrin, R. C., H. J. Nicholes and K. L. Siebecker. Am. J. Physiol. **123:** 98, 1938.
(194) Herrmann, G., G. Decherd and P. Erhard. Proc. Soc. Exper. Biol. and Med. **47:** 464, 1941.
(195) Herrmann, H., F. Jourdan and J. Vial. Compt. Rend. Soc. Biol., Paris **133:** 523, 1940.
(196) Hessel, G. Arch. f. exper. Path. **190:** 180, 1938.
(197) Hessel, G. Klin. Wchnschr. **17:** 843, 1938.
(198) Heymans, C. and J. J. Bouckaert. Proc. Soc. Exper. Biol. and Med. **39:** 94, 1938.
(199) Heymans, C., J. J. Bouckaert, L. Elaut, F. Bayless and A. Samaan. Compt. rend. Soc. de biol. **126:** 434, 1937.
(200) Hill, W. H. P. and E. C. Andrus. Proc. Soc. Exper. Biol. and Med. **44:** 213, 1940.
(201) Hill, J. R. and G. W. Pickering. Clin. Sci. **4:** 207, 1939.

(202) Hoffman, B. J. J. A. M. A. **120**: 1028, 1942.
(203) Holló, F. and Z. Kolbenheyer. Klin. Wchnschr. **19**: 302, 1940.
(204) Holman, D. V. and I. H. Page. Am. Heart J. **16**: 321, 1938.
(205) Holtz, P., R. Heise and K. Lüdtke. Arch. f. exper. Path. u. Pharmakol. **191**: 87, 1938.
(206) Horton, B. T. Proc. Mayo Clin. **15**: 472, 1940.
(207) House, R. M. and G. E. Wakerlin. Am. J. Physiol. **133**: 336, 1941.
(208) Houssay, B. A. and E. Braun-Menendez. Brit. med. J. **2**: 179, 1942.
(209) Houssay, B. A., E. Braun-Menendez and L. Dexter. Ann. Int. Med. **17**: 461, 1942.
(210) Houssay, B. A. and L. Dexter. Ann. Int. Med. **17**: 451, 1942.
(211) Houssay, B. A. and J. C. Fasciolo. Compt. rend. Soc. de Biol. **127**: 147, 1938.
(212) Houssay, B. A. and A. C. Taquini. Compt. rend. Soc. de Biol. **128**: 1125, 1938.
(213) Houssay, B. A. and A. C. Taquini. Compt. rend. Soc. de Biol. **129**: 860, 1938.
(214) Huchard, H. Traité clinique des maladies du coeur et de l'aorte, Paris, O. Doin, **1**: 45, 1899.
(215) Huidobro, F. and E. Braun-Menendez. Am. J. Physiol. **137**: 47, 1942.
(216) Johnson, C. A. and G. E. Wakerlin. Proc. Soc. Exper. Biol. and Med. **44**: 277, 1940.
(217) Johnson, C. A., G. E. Wakerlin and E. L. Smith. J. Immunol. **48**: 79, 1944.
(218) Jonnard, R. and M. R. Thompson. J. Am. Pharm. A. **32**: 260, 1943.
(219) Kahn, J. R. and T. C. Laipply. Am. J. Med. Sci. **203**: 807, 1942.
(220) Kaplan, A. and M. Friedman. J. Exper. Med. **76**: 307, 1942.
(221) Kapp, F., C. K. Friedland and E. M. Landis. Am. J. Physiol. **131**: 710, 1940.
(222) Katz, L. N. and L. Friedberg. Am. J. Physiol. **127**: 29, 1939.
(223) Katz, L. N., M. Friedman, S. Rodbard and W. Weinstein. Am. Heart. J. **17**: 334, 1939.
(224) Katz, L. N. and L. Leiter. Psychosom. Med. **1**: 101, 1939.
(225) Katz, L. N., M. Mendlowitz and M. Friedman. Proc. Soc. Exper. Biol. and Med. **37**: 722, 1938.
(226) Katz, L. N., S. Rodbard, F. S. Steinitz and L. Friedberg. J. Clin. Investigation **18**: 408, 1939.
(227) Katz, L. N. and F. S. Steinitz. Am. J. Physiol. **128**: 433, 1940.
(228) Katz, Y. J. and H. Goldblatt. J. Exper. Med. **78**: 67, 1943.
(229) Kaufmann, W. Am. J. Path. **18**: 783, 1942.
(230) Kempf, G. F. and I. H. Page. J. Lab. Clin. Med. **27**: 1192, 1942.
(231) Kempner, W. North Carolina M. J. **5**: 125, 1944.
(232) Kempner, W. North Carolina M. J. **5**: 273, 1944.
(233) Kohlstaedt, K. G. and I. H. Page. J. Exper. Med. **72**: 201, 1940.
(234) Kohlstaedt, K. G., I. H. Page and O. M. Helmer. Am. Heart J. **19**: 92, 1940.
(235) Lamport, H. Ann. Rev. Physiol. **7**: 331, 1945.
(236) Landis, E. M. Ann. Rev. Physiol. **2**: 141, 1940.
(237) Landis, E. M. Am. J. Med. Sci. **202**: 14, 1941.
(238) Lasher, E. P., Jr. and F. Glenn. Arch. Surg. **38**: 886, 1939.
(239) Landis, E. M. and W. A. Jeffers. J. Clin. Investigation **18**: 489, 1939.
(240) Landis, E. M., W. A. Jeffers and E. H. Shiels. Am. J. Physiol. **128**: 672, 1940.
(241) Leiter, L. and L. Eichelberger. J. Clin. Investigation **18**: 477, 1939.
(242) Leiter, L. and L. Eichelberger. J. Mt. Sinai Hosp. **8**: 744, 1942.
(243) Leiter, L. and L. Eichelberger. J. Clin. Investigation **22**: 11, 1943.
(244) Leloir, L. F., E. Braun-Menendez, L. Dexter, J. C. Fasciolo and A. Taquini. The N. Y. Acad. of Sci., Section of Biology, Feb. 9–10, 1945.
(245) Leloir, L. F., J. M. Muñoz, E. Braun-Menendez and J. C. Fasciolo. Compt. rend. Soc. de biol. **134**: 487, 1940.
(246) Leloir, L. F., J. M. Muñoz, E. Braun-Menendez and J. C. Fasciolo. Rev. de la Soc. Argent. de Biol. **16**: 635, 1940.

(247) Leloir, L. F., J. M. Muñoz, A. C. Taquini, E. Braun-Menendez and J. C. Fasciolo. Rev. Argent. Cardiol. **9:** 269, 1942.
(248) Leo, S. D., M. Prinzmetal and H. A. Lewis. Am. J. Physiol. **131:** 18, 1940.
(249) Levy, S. E., R. A. Light and A. Blalock. Am. J. Physiol. **122:** 38, 1938.
(250) Levy, S. E., M. F. Mason, T. R. Harrison and A. Blalock. Surgery **1:** 238, 1937.
(251) Levy, S. E., C. S. Robinson and A. Blalock. Am. J. Physiol. **123:** 383, 1938.
(252) Lewis, H. A. and H. Goldblatt. Bull. N. Y. Acad. Med. **18:** 459, 1942.
(253) Lewis, H. A., S. D. Leo and M. Prinzmetal. Am. Heart J. **21:** 319, 1941.
(254) Lisa, J. R., D. Eckstein and C. Solomon. Am. J. Med. Sci. **205:** 701, 1943.
(255) Lobo-Onell, C. and I. Diaz-Muñoz. Rev. med. latino-am. **26:** 1073, 1941.
(256) Longcope, W. T. and A. T. McClintock. Arch. Int. Med. **6:** 439, 1910.
(257) Loesch, J. Zentralbl. f. inn. Med. **54:** 145, 1933.
(258) Lorber, V. and M. B. Visscher. Am. J. Physiol. **133:** 365, 1941.
(259) MacLachlan, I. and N. B. Taylor. Am. J. Physiol. **129:** 413, 1940.
(260) Major, R. H. Arch. Int. Med. **62:** 946, 1938.
(261) Major, R. H., C. J. Weber and M. J. Rumald. Arch. Int. Med. **64:** 988, 1939.
(262) Mann, F. C., J. F. Herrick, H. E. Essex and E. J. Baldes. Surgery **4:** 249, 1938.
(263) Marienfeld, C. J. and G. E. Wakerlin. Fed. Proc. **2:** 32, 1943.
(264) Mason, M. F., C. S. Robinson and A. Blalock. J. Exper. Med. **72:** 289, 1940.
(265) Mason, M. F. and J. D. Rozzell. Proc. Soc. Exper. Biol. and Med. **42:** 142, 1939.
(266) McCann, W. S. and M. J. Romansky. J. A. M. A. **115:** 573, 1940.
(267) McEwen, E. G. and S. P. Harrison. Am. J. Physiol. **126:** 570, 1939.
(268) McEwen, E. G., S. P. Harrison and A. C. Ivy. Proc. Soc. Exper. Biol. and Med. **42:** 254, 1939.
(269) McManus, J. F. A. Lancet **2:** 394, 1942.
(270) McMartin, W. J. and T. McCurdy. J. Urol. **49:** 524, 1943.
(271) Megibow, R. S., L. N. Katz and S. Rodbard. Am. J. Med. Sci. **204:** 340, 1942.
(272) Merrill, A., R. H. Williams and T. R. Harrison. Am. J. Med. Sci. **196:** 240, 1938.
(273) Moritz, A. R. and M. R. Oldt. Am. J. Path. **13:** 679, 1937.
(274) Moschcowitz, E. Oxford Univ. Press, 1942.
(275) Moss, W. G., E. L. Smith and G. E. Wakerlin. Fed. Proc. **2:** 35, 1943.
(276) Moss, W. G. and G. E. Wakerlin. Proc. Cent. Soc. Clin. Res. **16:** 50, 1943.
(277) Muñoz, J. M., E. Braun-Menendez, J. C. Fasciolo and L. F. Leloir. Nature, London **144:** 980, 1939.
(278) Muñoz, J. M., E. Braun-Menendez, J. C. Fasciolo and L. F. Leloir. Am. J. Med. Sci. **200:** 608, 1940.
(279) No reference.
(280) Muñoz, J. M., A. C. Taquini, E. Braun-Menendez and J. C. Fasciolo. Rev. Soc. argent. de. biol. **19:** 321, 1943.
(281) Murphy, F. D., J. Grill, G. P. Langenfeld, L. J. Kurten and V. G. Guenther. J. A. M. A. **118:** 1245, 1942.
(282) No reference.
(283) No reference.
(284) Ogden, E. Texas Reports on Biol. Med. **2:** 345, 1944.
(285) Ogden, E., L. T. Brown and E. W. Page. Am. J. Physiol. **129:** 560, 1940.
(286) Ogden, E., E. W. Page and E. Anderson. Am. J. Physiol. **141:** 389, 1944.
(287) Oppenheimer, B. S., P. Klemperer and L. Moschkowitz. Trans. Assn. Am. Phys. **54:** 69, 1939.
(288) Oppenheimer, B. S., B. E. Lowenstein and C. Hyman. Proc. Soc. Exper. Biol. and Med. **57:** 117, 1944.
(289) Oster, K. A. and O. Martinez. J. Exper. Med. **78:** 477, 1943.
(290) Oster, K. A. and H. Sobotka. J. Pharmacol. **78:** 100, 1943.
(291) Oster, K. A. and S. Z. Sorkin. Proc. Soc. Exper. Biol. **51:** 67, 1942.

(292) Page, E. W., H. S. Patton and E. Ogden. Am. J. Obst. and Gynec. **41**: 53, 1941.
(293) Page, I. H. Am. J. Physiol. **112**: 166, 1935.
(294) Page, I. H. J. A. M. A. **113**: 2046, 1939.
(295) Page, I. H. J. Exper. Med. **70**: 521, 1939.
(296) Page, I. H. Science **89**: 273, 1939.
(297) Page, I. H. Am. Assn. Adv. Sc., no. 13, 239, 1940.
(298) Page, I. H. Am. J. Physiol. **130**: 22, 1940.
(299) Page, I. H. Am. J. Physiol. **130**: 29, 1940.
(300) Page, I. H. J. Exper. Med. **72**: 301, 1940.
(301) Page, I. H. J. Urol. **46**: 807, 1941.
(302) Page, I. H. Am. Heart J. **23**: 336, 1942.
(303) Page, I. H. J. A. M. A. **120**: 757, 1942.
(304) Page, I. H. Am. J. Physiol. **139**: 386, 1943.
(305) Page, I. H. Anesthesia and Analgesia **22**: 196, 1943.
(306) Page, I. H. Bull. N. Y. Acad. Med. **19**: 461, 1943.
(307) Page, I. H. J. Exper. Med. **78**: 41, 1943.
(308) Kohlstaedt, K. G., I. H. Page and O. M. Helmer. Am. Heart J. **19**: 92, 1940.
(309) Page, I. H. and O. M. Helmer. J. Exper. Med. **71**: 29, 1940.
(310) Page, I. H. and O. M. Helmer. J. Exper. Med. **71**: 495, 1940.
(311) Page, I. H., O. M. Helmer, K. G. Kohlstaedt, P. J. Fouts and G. F. Kempf. J. Exper. Med. **73**: 7, 1941.
(312) Page, I. H., O. M., Helmer, K. G. Kohlstaedt, P. J. Fouts, G. F. Kempf and A. C. Corcoran. Proc. Soc. Exper. Biol. and Med. **43**: 722, 1940.
(313) Page, I. H., O. M. Helmer, K. G. Kohlstaedt, G. F. Kempf, A. C. Corcoran and R. D. Taylor. Ann. Int. Med. **18**: 29, 1943.
(314) Page, I. H., B. McSwain, G. M. Knapp and W. D. Andrus. Am. J. Physiol. **135**: 214, 1941.
(315) Page, I. H. and J. E. Sweet. Am. J. Physiol. **120**: 238, 1937.
(316) Palmer, R. S., R. Chute, N. L. Crone and B. Castleman. New England J. Med. **223**: 165, 1940.
(317) Patton, H. S., E. W. Page and E. Ogden. Surg., Gynec. and Obst. **76**: 493, 1943.
(318) Perera, G. A. J. A. M. A. **129**: 537, 1945.
(319) Perera, G. A., A. I. Knowlton, A. Lowell and R. F. Loeb. J. A. M. A. **125**: 1030, 1944.
(320) Pickering, G. W. Clin. Sci. **2**: 209, 1936.
(321) Pickering, G. W. and M. Prinzmetal. Clin. Sci. **3**: 211, 1938.
(322) Pickering, G. W. and M. Prinzmetal. Clin. Sci. **3**: 357, 1938.
(323) Pickering, G. W. and M. Prinzmetal. Am. J. Physiol. **98**: 314, 1940.
(324) Pickering, G. W., M. Prinzmetal and A. R. Kelsall. Clin, Sci. **4**: 401, 1942.
(325) Plentl, A. A. and I. H. Page. J. Biol. Chem. **147**: 135, 1943.
(326) Plentl, A. A. and I. H. Page. J. Exper. Med. **79**: 205, 1944.
(327) Plentl, A. A., I. H. Page and W. W. Davis. J. Biol. Chem. **147**: 143, 1943.
(328) Plentl, A. A. and I. H. Page. J. Biol. Chem. **155**: 363, 1944.
(329) Plentl, A. A. and I. H. Page. J. Exper. Med. **78**: 367, 1943.
(330) Plentl, A. A. and I. H. Page. J. Exper. Med. **79**: 205, 1944.
(331) Plentl, A. A. and I. H. Page. J. Biol. Chem. **158**: 49, 1945.
(332) Prinzmetal, M., G. A. Alles, C. Margoles, S. Kayland and D. S. Davis. Proc. Soc. Exper. Biol. and Med. **50**: 288, 1942.
(333) Prinzmetal, M., B. Friedman and D. I. Abramson. Ann. Int. Med. **12**: 1604, 1939.
(334) Prinzmetal, M., B. Friedman and E. T. Oppenheimer. Proc. Soc. Exper. Biol. and Med. **38**: 493, 1938.
(335) Prinzmetal, M., H. A. Lewis and S. D. Leo. J. Exper. Med. **72**: 763, 1940.
(336) Prinzmetal, M., H. A. Lewis, J. Taggart, H. Wilkins and D. R. Drury. Am. Heart J. **20**: 525, 1940.

(337) PRINZMETAL, M. AND C. WILSON. J. Clin. Investigation **15:** 63, 1936.
(338) QUINBY, W. C., L. DEXTER, J. A. SANDMEYER AND F. W. HAYNES. J. Clin. Investigation **24:** 69, 1945.
(339) RASKA, S. B. J. Exper. Med. **78:** 75, 1943.
(340) RASKA, S. B. J. Exper. Med. **82:** 227, 1945.
(341) RATLIFF, R. K. AND K. B. CONGER. J. Urol. **48:** 136, 1942.
(341a) REED, R. K., L. A. SAPIRSTEIN, F. D. SOUTHARD, JR. AND E. OGDEN. Am. J. Physiol. **141:** 707, 1944.
(342) REMINGTON, J. W., W. D. COLLINGS, H. W. HAYS, W. M. PARKINS AND W. W. SWINGLE. Am. J. Physiol. **132:** 622, 1941.
(343) REMINGTON, J. W., W. D. COLLINGS, H. W. HAYS AND W. W. SWINGLE. Proc. Soc. Exper. Biol. and Med. **45:** 470, 1940.
(344) RIGGS, T. F. AND R. W. SATTERTHWAITE. J. Urol. **45:** 513, 1941.
(345) RINEHART, J. F., O. O. WILLIAMS AND W. S. CAPPELLER. Arch. Path. **32:** 169, 1941.
(346) RISKIND, L. A. AND H. H. GREENE. J. A. M. A. **119:** 1016, 1942.
(347) ROBBERS, H. AND O. WESTENHOEFFER. Ztschr. f.d. ges. Exper. Med. **105:** 180, 1939.
(348) RODBARD, S. Am. J. Physiol. **133:** 429, 1941.
(349) RODBARD, S. AND L. N. KATZ. J. Exper. Med. **73:** 357, 1941.
(350) ROGOFF, J. M. AND E. MARCUS. J. A. M. A. **110:** 2127, 1938.
(351) ROGOFF, J. M., E. N. NIXON AND G. N. STEWART. Proc. Soc. Exper. Biol. and Med. **41:** 57, 1939.
(352) ROSE, B. AND P. WEIL. Am. J. Physiol. **126:** 614, 1939.
(353) RYTAND, D. A. J. Clin. Investigation **17:** 391, 1938.
(354) SAPHIR, O. AND J. BALLINGER. Arch. Int. Med. **66:** 541, 1940.
(355) SAPIRSTEIN, L. A., E. OGDEN AND F. D. SOUTHARD. Proc. Soc. Exper. Biol. and Med. **48:** 505, 1941.
(356) SAPIRSTEIN, L. A., R. K. REED AND F. D. SOUTHARD, JR. J. Lab. and Clin. Med. **29:** 633, 1944.
(357) SCARFF, R. W. AND N. H. MARTIN. Brit. J. Exper. Path. **22:** 309, 1941.
(358) SCHALES, O. J. Am. Chem. Soc. **64:** 561, 1942.
(359) SCHALES, O. AND F. W. HAYNES. Proc. Soc. Exper. Biol. and Med. **47:** 315, 1941.
(360) SCHALES, O., M. HOLDEN AND S. S. SCHALES. Arch. Biochem. **2:** 67, 1943.
(361) SCHALES, O., E. A. STEAD AND J. V. WARREN. Am. J. M. Sc. **204:** 797, 1942.
(362) SCHROEDER, H. A. Proc. Soc. Exper. Biol. and Med. **44:** 172, 1940.
(363) SCHROEDER, H. A. Am. J. Med. Sc. **204:** 62, 1942.
(364) SCHROEDER, H. A. Am. J. Med. Sc. **204:** 734, 1942.
(365) SCHROEDER, H. A. Science **95:** 306, 1942.
(366) SCHROEDER, H. A. AND M. H. ADAMS. J. Exper. Med. **73:** 531, 1941.
(367) SCHROEDER, H. A., M. H. ADAMS AND A. E. COHN. J. Clin. Investigation **20:** 442, 1941.
(368) SCHROEDER, H. J. AND A. E. COHN. J. Clin. Investigation **17:** 515, 1938.
(369) SCHROEDER, H. A. AND J. M. STEELE. J. Exper. Med. **72:** 707, 1940.
(369a) SELKURT, E. E. Am. J. Physiol. **144:** 395, 1945.
(370) SELKURT, E. E. The Clin. Bull. of the School of Med., W. R. U. and its Assoc. Hosp. **9:** 87, 1945.
(371) SELYE, H. AND C. E. HALL. Am. Heart J. **27:** 338, 1944.
(372) SELYE, H., C. E. HALL AND E. M. ROWLEY. Canad. M. A. J. **49:** 88, 1943.
(373) SELYE, H. AND E. I. PENTZ. Can. M. A. J. **49:** 264, 1943.
(374) SELYE, H. AND E. M. ROWLEY. J. Urol. **51:** 439, 1944.
(375) SELYE, H. AND H. STONE. Proc. Soc. Exper. Biol. and Med. **52:** 190, 1943.
(376) SEMANS, J. H. Bull. Johns Hopkins Hosp. **75:** 184, 1944.
(377) SENSENBACH, W. Arch. Int. Med. **73:** 123, 1944.
(378) SHANNON, J. A. Am. Rev. Physiol. **4:** 297, 1942.
(379) SHRADER, J. C., J. M. YOUNG AND I. H. PAGE. Am. J. Med. Sc. **205:** 505, 1943.
(380) SHORR, E., B. W. ZWEIFACH AND R. F. FURCHGOTT. Science **102:** 489, 1945.

(381) Shure, N. M. Arch. Int. Med. **70:** 284, 1942.
(382) Smith, H. W. Harvey Lecture **35:** 166, 1939.
(383) Smith, H. W., W. Goldring and H. Chasis. J. Clin. Investigation **17:** 263, 1938.
(384) Smith, H. W., W. Goldring and H. Chasis. Bull. N. Y. Acad. Med. **19:** 449, 1943.
(385) Solandt, D. Y., R. Nassim and C. R. Cowan. Lancet **238:** 873, 1940.
(386) Soloway, S. and K. Oster. Proc. Soc. Exper. Biol. and Med. **50:** 108, 1942.
(387) Stead, E. A. and P. Kunkel. J. Clin. Investigation **19:** 25, 1940.
(388) Steele, J. M. Proc. Soc. Exper. Biol. and Med. **41:** 86, 1939.
(389) Steele, J. M. and A. E. Cohn. J. Clin. Investigation **17:** 514, 1938.
(390) Steiner, A., D. M. Weeks and A. L. Barach. Am. Heart J. **19:** 708, 1940.
(391) Steinitz, F. S. and L. N. Katz. Am. J. Physiol. **126:** 639, 1939.
(392) Stofer, B. E. and L. L. Kline. Arch. Path. **35:** 681, 1943.
(393) Stollowsky, G. Zentralbl. f. inn. Med. **61:** 513, 1940.
(394) Sweeney, J. S. and J. M. Page. Ann. Int. Med. **19:** 1013, 1943.
(395) Swingle, W. W., A. R. Taylor, W. D. Collings and H. W. Hays. Am. J. Physiol. **127:** 768, 1939.
(396) Taggart, J. and D. R. Drury. J. Exper. Med. **71:** 857, 1940.
(397) Talbott, J. H., B. Castleman, R. H. Smithwick, R. S. Melville and L. J. Pecora. J. Clin. Investigation **22:** 387, 1943.
(398) Taquini, A. C. Am. Heart J. **19:** 513, 1940.
(399) Taylor, H. E. Canad. M. A. J. **47:** 24, 1942.
(400) Taylor, R. D. and I. H. Page. Am. J. M. Sc. **205:** 66, 1943.
(401) Thompson, K. W. Physiol. Rev. **21:** 588, 1941.
(402) Tigerstedt, R. and P. G. Bergman. Skand. Arch. f. Physiol. **8:** 223, 1898.
(403) Turnoff, D. and L. G. Rowntree. Science **93:** 281, 1941.
(404) Verney, E. B. and M. Vogt. Quart. J. Exper. Physiol. **28:** 253, 1938.
(405) Verney, E. B. and M. Vogt. Quart. J. Exper. Physiol. **32:** 35, 1943.
(406) Victor, J. Proc. Soc. Exper. Biol. and Med. **60:** 332, 1945.
(407) Victor, J., A. Steiner and D. M. Weeks. Arch. Path. **29:** 728, 1940.
(408) Vogt, M. Quart. J. Exper. Physiol. **30:** 341, 1940.
(409) Volhard, F. and T. Fahr. Die Brightsche Nierenkrankheit. J. Springer, Berlin, 1914.
(410) Wakerlin, G. E. and G. R. Chobot. Am. J. Physiol. **126:** 646, 1939.
(411) Wakerlin, G. E. and G. R. Chobot. Proc. Soc. Exper. Biol. and Med. **40:** 331, 1939.
(412) Wakerlin, G. E. and W. Gaines. Am. J. Physiol. **130:** 568, 1940.
(413) Wakerlin, G. E. and C. A. Johnson. Am. J. Physiol. **129:** 188, 1940.
(414) Wakerlin, G. E. and C. A. Johnson. J. A. M. A. **117:** 416, 1941.
(415) Wakerlin, G. E. and C. A. Johnson. Proc. Soc. Exper. Biol. and Med. **46:** 101, 1941.
(416) Wakerlin, G. E., C. A. Johnson, B. Gomberg and M. L. Goldberg. Science **93:** 332, 1941.
(417) Wakerlin, G. E., C. A. Johnson, W. G. Moss and E. L. Smith. J. A. M. A. **122:** 60, 1943.
(418) Wakerlin, G. E., C. A. Johnson, E. L. Smith, B. Gomberg, J. R. Weir, W. G. Moss and M. L. Goldberg. Am. Heart J. **25:** 1, 1943.
(419) Wakerlin, G. E., C. A. Johnson, E. L. Smith, W. A. Moss and J. R. Weir. Am. J. Physiol. **137:** 515, 1942.
(420) Wakerlin, G. E. and W. G. Moss. Proc. Soc. Exper. Biol. and Med. **53:** 140, 1943.
(421) Wakerlin, G. E., W. G. Moss and E. L. Smith. Science **96:** 161, 1942.
(422) Wakerlin, G. E. and M. R. Salk. Am. J. Physiol. **133:** 479, 1941.
(423) Wakerlin, G. E. and M. Yanowitz. Proc. Soc. Exper. Biol. and Med. **41:** 51, 1939.
(424) Weber, C. J., R. H. Major and D. Lobb. Science **96:** 44, 1942.
(425) Weeks, D. M., A. Steiner, J. S. Mansfield and J. Victor. J. Exper. Med. **72:** 345, 1940.

(426) WEINSTEIN, H., M. FRIEDMAN, H. L. NEWMANN AND J. SUGARMAN. Am. Heart J. **25**: 682, 1943.

(427) WEISS, E. AND H. CHASIS. J. A. M. A. **123**: 277, 1943.

(428) WEISS, S. AND F. PARKER, JR. Medicine **18**: 221, 1939.

(429) WHITE, B. V., R. E. DURKEE AND C. MIRABILE. New England J. Med. **228**: 277, 1943.

(430) WILLIAMS, J. R., JR. Am. J. Physiol. **124**: 83, 1938.

(431) WILLIAMS, J. R. JR., J. T. DIAZ, J. C. BURCH AND T. R. HARRISON. Am. J. Med. Sc. **198**: 212, 1939.

(432) WILLIAMS, J. R., JR., A. GROLLMAN AND T. R. HARRISON. Am. J. Physiol. **130**: 496, 1940.

(433) WILLIAMS, J. R., JR., A. GROLLMAN AND T. R. HARRISON. Arch. Int. Med. **67**: 895, 1941.

(434) WILLIAMS, J. R., JR., T. R. HARRISON AND M. F. MASON. Am. J. M. Sc. **195**: 339, 1938.

(435) WILLIAMS, J. R., JR., R. WEGRIA AND T. R. HARRISON. Arch. Int. Med. **62**: 805, 1938.

(436) WILSON, C. Proc. Roy. Soc. Med. **39**: 61, 1945.

(437) WILSON, C. AND F. B. BYROM. Lancet **1**: 136, 1939.

(438) WILSON, C. AND F. B. BYROM. Quart. J. Med. **10**: 65, 1941.

(439) WILSON, C. AND G. W. PICKERING. Clin. Sc. **3**: 343, 1938.

(440) WINTERNITZ, M. C., E. MYLON AND R. KATZENSTEIN. Yale J. Biol. and Med. **13**: 789, 1941.

(441) WINTERNITZ, M. C., E. MYLON, L. L. WATERS AND R. KATZENSTEIN. Yale J. Biol. and Med. **12**: 623, 1940.

(442) WINTERNITZ, M. C. AND L. L. WATERS. Yale J. Biol. and Med. **12**: 451, 1940.

(443) WINTERNITZ, M. C. AND L. L. WATERS. Yale J. Biol. and Med. **12**: 705, 1940.

(444) WOODS, W. W. J. Urol. **48**: 16, 1942.

(445) WOSIKA, P. H., F. T. JUNG AND C. C. MAHER. Am. Heart J. **24**: 483, 1942.

(446) YUILE, C. L. Am. J. Med. Sc. **207**: 394, 1944.

PHYSIOLOGICAL REVIEWS

VOL. 27 APRIL, 1947 No. 2

ON THE NATURE OF PAIN

HAROLD G. WOLFF AND JAMES D. HARDY

The New York Hospital and the Departments of Medicine (Neurology), Psychiatry and Physiology, Cornell University Medical College, New York, N. Y.

Concepts Concerning Pain. Views about the nature of pain have varied, but from Aristotle to relatively recent times there has been general agreement on one major issue, namely, that pain is a "passion of the soul", a feeling state, or a "quale", and not a specific sensation. Such a view was supported by thoughtful workers in nineteenth century Germany, England and America (1).

As recently as the turn of the twentieth century it was still considered debatable whether pain is indeed a sensation or exclusively a feeling reaction akin, but opposite, to that of pleasure. The following is representative of the prevailing view of that period: "... that pain is not a sensation but a form of feeling; that it is not to be classed with the sensations of touch, or temperature, or heat; that it does not have peripheral end organs, and that there are no nerves in existence which, on irritation alone, produce pain; that there is no such thing as a pain tract. . . ." (2).

However, there has been a gradual change in this conception and the basis for modifying this earlier view is briefly summarized in the following paragraphs.

1. Blix (3), Goldscheider (4) and von Frey (5) noted that when the skin is explored with stiff hairs, pain alone may be elicited from certain areas. The structures stimulated have been indentified by Woollard, Weddell and Harpmann (6) who demonstrated that cutaneous pain is subserved by the finer medullated and non-medullated nerve fibres bearing free endings. These fibres and endings are specific and are arranged in a plexiform interlocking manner. Weddell (7, 8) in order to determine the neurohistology of skin from which pain alone could be evoked, outlined the pattern of sensory loss in a patient with a sciatic nerve lesion. The area from which pain could not be aroused was smaller than that from which touch could not be aroused, and this in turn was smaller than the areas over which temperature sensibility had been lost. A piece of skin 2 by 3 cm. was stained and removed from a zone in part of which pain only could be aroused and in part of which pain and touch could be evoked. In the area from which pain alone could be elicited, fine nerve fibres were seen giving rise to superficial nerve nets; no thick nerve fibres or organized endings were seen. On the other hand, in the area of skin in which both pain and touch could be evoked, the cutaneous nerve plexus was seen, together with thick nerve fibres ending around hairs and thinner fibres giving rise to superficial nerve nets.

Tower (9) introduced the useful concept of the "sensory unit" as contrasted with the single afferent nerve ending. She has made it seem likely that a "sensory unit" consists of many nerve endings, all branches of a single fibre,

connected with a single cell in the dorsal root ganglion. An area of skin (12) or cornea (9) of about a centimeter in diameter is said to be supplied by such a single unit and it is conceivable that larger units exist. Many pain fibres supply overlapping "sensory units" to a given area, although there is no connection between the endings of different fibre units. It would thus seem likely that the painful spot experienced as such when an area of normal skin is exposed to noxious stimulation is a centrally integrated experience projected onto the periphery for the specific purpose of localization. The "skin spot" is, so to speak, a "mind spot".

2. An answer to the question as to whether pain is a specific sensation or whether it results from overstimulation of fibres serving other sensations, such as touch or warmth, is the demonstration in patients (10), that pain alone and no other sensation may be lost in limited regions of the body as a result of accidental injury of the spinal cord. Also, after surgical transection of the spinothalamic pathways, pain alone may be absent in circumscribed regions without associated loss of temperature or touch (11).

Moreover, it has been demonstrated by Adrian, Cattell and Hoagland (12) and by Cattell and Hoagland (13) that intensive stimulation of tactile end organs up to the capacity of the corresponding nerve fibres to conduct does not elicit evidences of pain.

3. Further indication that pain is a specific sensory experience with its own neural structures and properties has been afforded by the analysis of the action of certain analgesic drugs. These agents were capable of specifically raising the threshold of pain from 35 to 80 per cent above the control level while they lowered or left unaltered the threshold for the perception of touch, hearing, smell, two point discrimination and the perception of vibration. It has been observed that acetylsalicylic acid actually lowers the threshold for warmth (14, 15, 16, 17).

4. There is evidence to indicate that certain structures are equipped for and give rise to the sensation of pain only. The teeth, the middle meningeal artery and the arteries at the base of the brain and at least some of the scalp arteries, such as the temporal artery, seem to be equipped with afferent fibres which, as far as consciousness is concerned, conduct only those impulses which result in pain (18). However, at least one other sensation, i.e., cold, and possibly touch as well as pain can be elicited by suitable stimulation of the cornea (19).

On the other hand, there are certain areas in the body from which pain cannot be elicited, for example, the parenchyma of the brain (18). Also, such an area is found on the inside of the cheek opposite the second lower molar (20). In short, it is necessary to infer from the evidence that the pain experience is first a sensation derived from noxious impulses traversing specific pathways. Such phenomena may be followed by the familiar and predictable feeling states and other reactions.

Neural Structures Involved in the Conduction of Noxious Impulses Giving Rise to Pain. All fibres carrying noxious impulses experienced as pain enter the spinal cord or brain stem through the dorsal root ganglia. The superficial

pain impulses are usually conveyed by the somatic nerves and enter the cord more or less directly through the dorsal roots. Deep pain impulses may approach the central nervous system in a number of ways: some are conveyed by pain fibres which attach themselves to blood vessels part of the way and then join autonomic nerves; other deep pain fibres from their very beginnings are closely associated with autonomic nerves and remain so affiliated until relatively near the cord; and still other deep pain fibres join with somatic nerves. The course they assume in approaching the dorsal root ganglia is irrelevant to a consideration of the quality of pain or the reflexes aroused (21, 22, 23, 24, 25, 26, 27, 28).

Although pain endings are usually naked and umyelinated, with or without slightly swollen twigs, there is no uniformity about the size of fibres which conduct noxious impulses to the cord and brain. Fibres conducting pain impulses may be roughly grouped as regards speed of conduction of impulses into slow and rapid, involving respectively small and large calibre fibres (29, 30). It has been shown by Gasser (30) that the gamut of size involved in the conduction of noxious impulses giving rise to pain is wide, spreading from the very smallest to large calibre fibres.

After entering the cord all noxious impulses are conveyed across to the opposite side where their pathways are localized in the anterolateral portion of the spinal cord. The fibres of the spinothalamic tract pass into the nucleus centralis posterior of the thalamus. They do not terminate in any of the adjacent nuclei nor go into the anterior portion of the thalamus. The cortical projection from the nucleus centralis posterior is predominantly to the post-central convolution. There is in this projection a definite topical organization so that those fibres from the medial portion of the nucleus (cephalad parts of the body) end in the lower part of the gyrus; those from the lateral portion (caudad parts of the body) in the paracentral region; and those from the middle, in the intermediate region.

It is probable that the brain structures involved in pain perception occur in both cerebral hemispheres in the region of the central fissure (31, 31a, 32). Complete hemianalgesia is rare after destruction of fibres in either hemisphere. More specifically, pain from the face seems to have both homolateral and contralateral cortical representation so that unilateral hemispherectomy does not result in analgesia of the face. Although the thalamus is important in the integration of neural activity having to do with pain, the cortex is essential to discriminations involving localization and intensity. Excision of a small portion of the left post-central cortex has resulted in elimination of phantom limb pain for a period exceeding 4 years (32a). That there is also mid-brain sensory integration is suggested by the survival of pain sensation after destruction of the fibres entering the thalamus and cortex.

Qualities of Pain and Their Significance. For purpose of gross classification superficial and deep pain are recognized. Pain resulting from noxious impulses originating on the surface or superficially has a pricking, bright, burning, itching quality, and is highly localizable. It commonly incites action such as fight or flight. Pain resulting from deep noxious impulses is a deeper, diffuse pain,

has an aching quality and localization is less precise. It often induces nausea, sometimes fall in blood pressure, prostration and syncope and usually causes the animal to seek protection in withdrawal, inactivity and rest. There are minor differences in the deep pain from muscle, web and artery. Structures of endodermal origin are said to be relatively poorly supplied with pain fibres as compared with those of mesodermal or ectodermal origin (26).

Man's ability to experience pain is not essential to a suitable biological adjustment. Persons congenitally without the ability to experience pain adequately adjust themselves to their environment, as do persons who have had pain pathways surgically interrupted (33, 34). Alarm or defense reactions may be initiated by any stimulus if that stimulus has been previously associated with injuries, dangerous threat situations or frustrations. As a matter of fact, the bulk of such reactions involved in common experience are initiated by non-painful stimuli. Pain is an accessory though important and final warning of the imminence of tissue damage.

On the other hand, persons critically ill, as with terminal neoplastic disease, and those gravely injured do not inevitably experience intense pain. Indeed, the intensity of pain is not directly proportional to the extent or seriousness of tissue damage. In a combat zone during World War II Beecher (35) questioned 215 recently wounded men concerning the intensity of the pain they were experiencing. These men had sustained extensive soft tissue injuries, compound fractures or penetrating wounds of the head, chest or abdomen, but were mentally clear. Only twenty-four per cent had "bad pain" whereas the remainder had moderate, slight or no pain.

Pain Threshold. Pain is best investigated in conscious man, otherwise its study resolves itself into an analysis of reactions. Hence, at the outset a fundamental distinction must be made, namely, between the perception of pain and its associated reactions. A starting point for a quantitative analysis of pain perception can be made by defining the pain threshold and by ascertaining its properties, and how it can be modified (36, 37).

The pain threshold sensation may be defined as the lowest perceptible intensity of pain. The pain threshold stimulus is that amount of stimulus required to induce threshold pain and may for convenience be expressed in standard physical terms. In accordance with common usage the pain threshold is considered to be raised when more stimulus is required to induce threshold pain; conversely, when less stimulus is required the pain threshold is said to be lowered. Pain is unique in that many methods of stimulation can evoke it. Both superficial and deep pain can be elicited by thermal, electrical, mechanical and chemical stimuli. It is only necessary to choose a stimulus, the strength of which can be controlled and measured, which will allow a clearly definable end point of perception, and one which will not of itself introduce excessive variability into the observation.

For the establishment of the threshold of pain the verbal report of an instructed subject is the most reliable evidence. A verbal report of an end-point has been successfully used in the researches on other sensations and in

studies on cutaneous pain has yielded reproducible data. Muscle twitch, blinking, withdrawal or any similar response indicating reaction to pain should be noted together with, but independently of, the verbal report of the pain threshold of a human being.

1. *Methods.* Ivy and his colleagues have considered all the methods that have been used for measuring the pain threshold and have listed a complete bibliography (38). Many methods for ascertaining the pain threshold have produced data which have been difficult of interpretation. This is possibly because of a) the failure to establish a known relationship between the quantity measured as the stimulus and the amount of pain producing disturbance in the environment of the pain endings; and b) the fact that reaction to a noxious stimulus is often confused with perception of threshold pain.

For example, in electrical stimulation of the teeth the amperage, voltage, frequency and resistance in the circuit are all of importance in inducing pain. Measurement of one of these quantities as the pain threshold might be expected to lead, because of its indirect relationship to the amount of disturbance in the environment of the pain endings, to considerable variability. Indeed, this has been found to be the case for this method of investigating pain threshold by Ivy et al, by Roth and Kleitman and by Robertson, Goodell and Wolff (38, 39, 40, 41). These same general criticisms also hold for the mechanical methods for eliciting pain, such as measuring the pressure required to evoke pain in the esophagus, the gut, or in the skin (42).

Chapman and Jones (42) measured the pressure required to induce pain in the esophagus by distention, and Harrison and Bigelow (43) also studying deep pain measured the point at which pain occurred in working ischemic muscles of the forearm.

Other investigators, notably Libman (44) and later Hollander (45) whose investigations were directed toward estimation of the threshold for pain perception actually tested reactivity. Libman exerted pressure with his thumb upon the styloid processes of human subjects and found that prize fighters, Negroes and American Indians as groups failed to react to such noxious stimulation of intensity great enough to induce a reaction of discomfort in the average white city dweller.

Hollander's instrument was a rough, metal grater incorporated in a sphygmomanometer cuff. He inflated the cuff and noted the pressure at which the subject winced. He, as did Libman, reported that relatively stoical people have a high threshold for reacting to noxious stimulation.

In spite of the above mentioned quantitative and conceptual inadequacies of many methods for studying pain threshold they have contributed much to the general body of information concerning the pain experience. Any serious consideration of the nature of pain sensation must take full account of the data derived by these methods.

Bazett and his colleagues have demonstrated that all sensation elicited by heating the surface of the skin is the direct result of the change in the thermal gradients near the skin surface (46). Also, it has been demonstrated that the

change in the thermal gradients in the skin is directly proportional to the intensity of the thermal energy brought to bear upon the skin (47), thereby demonstrating that this quantity, measured as the stimilus, is directly proportional to the algesic, that is, the amount of pain producing disturbance at the pain ending. This stimulus can be expressed quantitatively in gram calories/second/cm.[2] Therefore, to measure the threshold of cutaneous pain the thermal irradiation of the skin has proved to be satisfactory and by this method both pricking and a burning pain threshold can be ascertained (37, 48). For most experimental purposes, a pricking pain threshold is more convenient. Most of the observations on cutaneous pain considered below are in terms of pricking pain threshold.

2. *Uniformity of the Cutaneous Pain Threshold in Man.* It was observed that successive observations of the pricking pain threshold on the forehead of a single individual could be reproduced within ±3 per cent (thermal irradiation method). The pain thresholds for a group of 200 individuals were thus measured (49). They were found to require the same intensity of stimulus to evoke pain, that is, 0.218 gram calories/second/cm.[2] with a standard deviation of ±5 per cent, maximum variation of ±15 per cent. The group of 200 included individuals of both sexes from ten to eighty years of age.

In order to ascertain whether pain threshold varies with the emotional states of the subjects, measurements were made in three individuals at the same time each day over a period of several months. Immediately preceding the measurement each day, the subjects recorded a statement of their estimated general effectiveness and mood. Although during this period these estimates showed wide variation, the pain threshold measured in these subjects did not vary beyond normal limits nor consistently with the general effectiveness or the emotional state. Also, throughout a twenty-four hour period of enforced wakefulness the pain thresholds were uniform and normal (49).

From these data it is inferred that the pain threshold in man is relatively stable and uniform when measured in the manner described. Furthermore, it is independent of age, sex, various emotional states and fatigue, and the time of day within the limits of the ability of the subject to maintain proper attention (37, 49).

Factors Which Influence Measurement of Cutaneous Pain Threshold in Man. Having established the conditions for the uniformity of the pain threshold, the factors that alter it may be considered. These may be roughly classified as physical, physiological, chemical and psychological.

The several physical factors which were observed to affect the pain threshold are those that vary the rate of heat loss from the skin. Also, the thickness of the skin (calluses) and the skin temperature, alter the amount of energy required to elicit pain. Variations in the intensity of stimulus required to elicit threshold pain on different parts of the same body may be accounted for by these factors. To eliminate these physical factors most tests were made on the forehead because the skin temperature of the forehead and the thickness of its skin are uniform from person to person. Starting with the periphery

it is evident that two effects may be expected from the peripheral endings responsible for pain. The threshold may be lowered by traumatic deformation, denudement or injury of tissues near these endings. There is experimental evidence that the pain threshold may be lowered as much as fifty per cent in sunburned skin (50). Other experiments show that by local anesthetization the threshold can be greatly raised or even obliterated. Obviously, blocking or cutting through a nerve fibre prevents it from carrying any impulses.

When a peripheral nerve is slightly injured, as by sustained pressure, a pin prick administered with constant force may feel less sharp, although the pricking pain threshold in the area of skin supplied by that nerve may be normal. However, if the nerve injury be sufficiently great, not only do pin pricks feel less sharp, but the pricking pain threshold, as ascertained by the radiation technique, is raised.

A raised pain threshold was observed in hypoalgesic areas of skin in patients with syringomyelia and damage of the spinothalamic tracts. A stimulus strong enough to cause blistering was sometimes required to evoke pain. In patients with lesions in the region of the internal capsule, elevated pain thresholds were also observed in the involved areas of skin, where pin prick felt less sharp. Furthermore, in patients with so-called central pain and lesions near the thalamus, with burning sensation and other paresthesias associated with the complaint that pin prick felt sharper than normal, the pain threshold in the involved painful area was usually elevated and never depressed (37, 48).

Also, there are other factors that may change the level of the pain threshold (37). Distraction, suggestion and hypnosis have been demonstrated to raise the pain threshold as much as thirty-five per cent. It is also observed that certain persons, during periods of fatigue, lethargy, suggestibility, prejudice or anxiety were unable to focus on the procedure or to maintain an objective attitude toward the sensation. In such persons the pain threshold varied greatly and was unpredictable. In the assay of analgesic agents, the attitudes and attention of the subject became especially important. Entirely different measurements of pain threshold were obtained depending on the ability of the subject to give attention to the assay. In a suggestible or prejudiced subject placebos significantly raised the pain threshold, whereas ordinarily effective amounts of analgesics produced minimal pain threshold raising effect (51).

It has long been known that "of two pains occurring together, not in the same part of the body, the stronger weakens the other" (Hippocrates). A corollary of this ancient observation is the experimental demonstration of a raised pain threshold in persons experiencing pain spontaneously or with pain experimentally induced. Gammon and Starr (52), although they were concerned with the effects on intensity of sustained pain rather than on pain threshold made similar observations. According to these authors the intensity of sustained deep pain induced by the subcutaneous injection of ten per cent sodium chloride into the arm is appreciably, though temporarily, reduced by stimulation of the skin adjacent to the area of pain. The intensity of pain was modified not only immediately after the application of the counter-irritant, but again

immediately after it had been withdrawn. In pain due to ointments applied on the skin, only cold applied to the painful areas reduced its intensity.

Effects of Analgesics on Pain Perception. Predictable effects as regards elevation of pain threshold were obtained with a variety of analgesic agents (14, 15, 17, 53, 54, 55, 56, 57, 58, 59). Also, maximal analgesic effects were achieved by given amounts of a drug and larger amounts within pharmacological limits produced no further threshold raising effect (14, 15, 59). However, as one approached narcosis following very large amounts of an analgesic agent, further threshold elevation was achieved (59). With smaller amounts, the effect on the pain threshold was proportional to the amount. Each of the drugs studied had its own maximal or ceiling amount which produced the maximal threshold raising effect. Furthermore, analgesics acted on pain with approximately the same effectiveness whether the noxious stimuli were conducted by smaller or by larger fibres (48, 59).

To ascertain whether these generalizations concerning threshold-raising effects are valid for deep pain as well as for cutaneous pain, Harrison and Bigelow (43) used a modification of Lewis' method for inducing deep pain (26). By repeatedly clenching the fist at the rate of once per second, after establishing asphyxial block by means of a ligature about the upper arm, a predictable number of contractions could be established as the threshold for pain. It was observed that the number of contractions of the hand necessary to induce pain was increased after analgesics. In terms of percentage, the effect approximated that for the cutaneous pain threshold obtained by the radiation technique. Thus, the aforementioned generalizations concerning the action of analgesics are valid for both cutaneous and deep pain. Moreover, there exists no satisfactory evidence to support the view that one or another analgesic agent is dependent for its effects upon the type or the source of pain.

Experiments were performed to ascertain whether there be a quantitative relationship between the change in pain threshold resulting from the action of an analgesic and the intensity of pain produced by standard stimuli well above the threshold (60) (see next section). It was observed that after the administration of a threshold-raising agent, the estimation of experimentally induced pain is strikingly modified. Thus, when the pain threshold is raised approximately forty per cent, stimuli which previously induced pain of three plus intensity are no longer painful, and those stimuli that normally induce pain of seven to eight plus intensity induce pain of three to four plus intensity.

Quantitative study of the effect of an analgesic agent in reducing the intensity of the spontaneous pain of disease is just beginning. This important question has not been approached hitherto because of lack of experimental methods, and because deductions regarding analgesic action have been made largely from the agent's effects upon the pain threshold of normal subjects. Some of the difficulties of interpreting these data in terms of the individual in pain have been pointed out above.

It has long been known that pain is a powerful antagonist of morphine action, and it has been demonstrated that prolonged pain diminishes or obliterates

the threshold-raising action of morphine (14). Depending on the time the painful experience was experimentally introduced in terms of the curve of threshold-raising effect of the morphine, a greater or lesser degree of disturbance in threshold-raising action was observed. Thus, if a pain of a given intensity and duration was experienced in the course of the action of an analgesic agent, it modified the threshold-raising effect but little. If it was experienced early in the course, a gross reduction of threshold-raising effect was observed. Despite the fact that there was relatively little elevation of pain threshold under the latter circumstances, the subjects seemed to be indifferent to the occurrence of pain. Although pain may be recognized as such, the individual seems to lose his anxiety and other reactions that usually accompany pain perception. It would appear, then, that with morphine one can demonstrate the distinction between pain perception and pain reaction.

The Discrimination of Intensity of Pain. A new approach to the study of the pain experience has been made in attempts to evaluate the intensity of pain above the pain threshold and to correlate this evaluation with the intensity of the evoking stimulus. Experiments of this nature on other sensations, particularly vision, hearing and heat, have been carried out and have added much to the understanding of the sensory mechanism involved (60, 61, 62, 63, 64, 65).

The ability to discriminate small differences in intensity of painful stimuli has been investigated by means of the thermal radiation method (60). This discrimination was studied in the intensity range from the pain threshold (218 millicalories/sec./cm.^{2}) to intensities of more than twice this value (480 millicalories/sec./cm.^{2}). It was found that from the pain threshold to forty-five per cent above, it was possible to distinguish differences in intensity of painful stimuli of ± 4 per cent of the threshold stimulus. As the intensity of painful stimulus was increased beyond 345 millicalories/sec./cm.^{2} this keen discrimination gradually diminished. With intensities of 480 millicalories/sec./cm.^{2}, i.e., more than twice that of the pain threshold, differentiation for higher intensities was no longer possible. Between threshold stimulus and the stimulus beyond which no further discrimination was possible (i.e., the highest intensity of pain which can be perceived), approximately twenty-one just noticeable differences of pain intensity could be distinguished. With the higher intensities of stimulus tissue damage occurred.

In this connection it is important to point out the fine discrimination possible for pain sensation, i.e., twenty-one discriminable steps for a two-fold increase in stimulus intensity. In contrast, the range of energy required to elicit a skin sensation by thermal stimulation beginning at the threshold of temperature or warmth sensation and ascending to the pain threshold is as one is to 2000 (63). There appear to be only about ninety steps of intensity of sensation distinguishable between warmth threshold and the onset of pain. Also, the energy range for vision from the threshold to the dazzle point is as one is to ten billion and there are only 572 distinguishable steps in this enormous range (66).

Certain individuals commonly overestimate the intensity of spontaneously

occurring or experimentally induced pain, but many of these individuals are able, when given a standard painful stimulus, to distinguish various levels of intensity. Although there are some persons who are untrustworthy witnesses of pain intensity, they are identified by their overreaction to non-noxious stimuli and their anxiety concerning or prejudice toward the experimental procedure.

Evidence Concerning Spatial Summation for Cutaneous Pain. There are two aspects of spatial summation. The first which has been reported for vision (61) and for heat sense (62) is that the intensity of sensation from a fixed intensity of stimulus increases or decreases with decrease or increase in the size of the area stimulated. Although experience indicates that noxious impulses may summate under certain circumstances there is at present no quantitative evidence on this point for pain sensation. The second aspect of spatial summation which has been demonstrated for vision (61) and for cutaneous temperature sensation (62) is that the sensory threshold is decreased or increased as the size of the area stimulated is increased or decreased. The usual explanation of this effect is that the sub-threshold impulses from separate end organs are summated in the central nervous system to produce a sensation (35). The significance of spatial summation to the body economy is that it provides extreme sensitivity, thereby permitting sensation to be evoked by minute changes in the environment. For example, vision is stimulated by as little as fifty-eight quanta of energy when the entire eye is illuminated, and warmth sense is evoked by a rise in skin temperature of 0.0008° C/sec. when the entire body surface is stimulated (62).

That spatial summation does not occur when the number of noxious impulses originating in the periphery is greatly increased above the control pain threshold level was demonstrated by the fact that morphine had the same threshold raising effect whether measured in small or in large areas of skin.

For the study of cutaneous pain threshold it was of importance to ascertain whether pain sense exhibited the phenomena of spatial summation. It was observed that within the limits of application of the method, the pain threshold for normally innervated skin was dependent alone upon the strength of the stimulus and not upon the size of the area stimulated, i.e., it was independent of the number of end organs stimulated (36).

Measurements of pain threshold by the radiation technique for areas of skin 0.3 cm.^{2} in size and smaller show an increase in the pain threshold as the area is decreased. This would appear to be evidence of spatial summation of cutaneous pain but careful measurements have shown that due to diffraction and thermal conduction in the skin, the increase in skin temperature produced in such small areas by a specified intensity of radiation is not as great as in larger areas. That is, the limit for comparison of area for the radiation technique is approximately 0.3 cm.^{2} and the apparent spatial summation is actually a measure of the progressive failure of the radiation to raise the skin temperature in such small skin areas. This artefact demonstrates the importance of establishing a known relationship between the quantity measured as the stimulus and its pain evoking action.

Another apparent exception to the concept of non-spatial summation is pre-

sented by subjects who had the number of pain endings in the tissue studied greatly reduced as a result of sensory root section. Thus, in the case of a patient who had his seventh and eighth cervical roots sectioned, the pain threshold was average and uniform on the lateral or normally innervated side of the hand. It was distinctly elevated on the medial portion or that supplied by the sectioned roots. The pain threshold rose abruptly as it approached the most medial portion of the hand and fingers. It is conceivable that the elevated pain threshold in the margin of the nerve-free region is the result of an inadequate density of pain endings in the tissue. It is, however, more likely that the perception of any pain in this area was due to a defect in the method of study, in that the strong thermal stimulus was actually stimulating distant pain endings in the intact portion of the hand through poorly conducting skin tissue.

The importance of the lack of spatial summation is twofold: first, it distinguishes cutaneous pain physiologically from cutaneous temperature sense; secondly, it implies that the intensity of a pain is dependent upon the intensity of the stimulus and not upon the size of the area involved, although the total distress to the individual is dependent both upon the intensity of the stimulus and the size of the painful area and other factors (36, 60).

Moreover, associated with this lack of spatial summation for pain the range of stimuli for producing pain up to the point of tissue damage is small (60). It has been repeatedly demonstrated on the skin that a stimulus twice the pain threshold value causes tissue destruction. In contrast with this is the wide range of stimuli between the threshold for warmth and the pain threshold. It was mentioned that the range of stimulus energy from the excitation threshold of warmth or temperature sensations up to the pain threshold is as 2000 is to one (62, 63). Here spatial summation is of the utmost importance for orientation, and permits a wide margin for discriminative actions of aggression and defence. Through pain the organism is warned of the imminence of tissue damage. There is no need in the body economy for a wide range of experience between the onset of pain and tissue damage since with the first pain impulses the organism receives warning that the limit of safety has already been passed. On the other hand, lack of spatial summation of pain provides that the organism will not be overwhelmed by noxious impulses and so prevented from effective action.

Two Types of Cutaneous Pain. The earliest mention of the view that there are two types of cutaneous pain was made by Gad and Goldsheider in 1892 (67). Since then the double nature of cutaneous pain has been repeatedly explored and although investigators differ concerning whether or not there are two qualities of cutaneous pain or merely variations of one, all are agreed as to the difference in speed of perception. Woollard, Weddell and Harpman (6), Boring (68) and more recently Bigelow, Harrison, Goodell and Wolff (48) are of the opinion that two qualities of pain are involved. Lewis (26) and Gasser (30) are less convinced of differences in quality of sensation but find the difference in the speed of perception of special interest. According to Bigelow, Harrison, Goodell and Wolff (48), one type of pain is abrupt in onset and has a pricking quality which quickly terminates; the other type is slow in onset, never reaches

an abrupt or pricking climax and more slowly recedes. The latter pain has a burning quality whether initiated by pin prick or by heat.

Bishop (69, 70, 71, 72) suggested that "itch, nonpainful prick, and pain, elicited by appropriate patterns of stimulation from the same point differ in quality as well as quantitatively." This qualitative shift with a changing quantity of stimulation indicates, according to Bishop, "a central qualitative interpretation of sensory impulses depending on quantitative factors involving identical peripheral mechanisms within the single modality of pricking pain". It must be pointed out, however, that "stimulation from the same point" is not equivalent to stimulating the same single nerve ending, and therefore does not ensure the implication of an "identical peripheral mechanism" for touch, itch, prick and pain.

Gasser's (73, 74) experiments indicate that the pain first to be perceived when a pin is pressed into the skin or when a finger touches a very hot electric bulb, the so-called "first pain", is carried mainly by the larger medullated fibres conducting at a rate between 10 and 90 metres per second, whereas the slow pain, or that to come on after the above pricking pain has been experienced (so-called "second pain") is carried mainly by small fibres called C fibres at a rate of between 0.6 and 2 metres per second.

Bigelow, Harrison, Goodell and Wolff (48) using the radiation technique studied the difference in properties of these two types of pain. Instead of "pricking" pain endpoint with its final abrupt, sharp-pointed quality, they measured and used the sensory threshold of "burning" pain which was thirteen per cent lower than the threshold for pricking pain for a 5 second exposure to radiation. It had a painful quality which was easily recognized. They showed that this second or burning pain sensation had many properties in common with pricking pain in that the threshold was raised after administering alcohol or morphine, and that no spatial summation occurred when larger and larger areas were stimulated.

Hence, under the circumstances mentioned, there was studied the threshold of another type of cutaneous pain akin, perhaps, to Lewis' "second pain" (26, 75) and conveyed probably by Gasser's "C" fibres (73, 74). The forehead was examined from the point of view of separating these two thresholds and it was noted that over a considerable range in the time of exposure, the pricking pain threshold was appreciably higher than the burning pain threshold. The difference between the two thresholds is more difficult to establish when the tissue is briefly stimulated.

From observations of these two types of pain it became possible to understand certain pathological states. It appeared from studies of pricking pain threshold that there was no lowering of pain threshold in patients with hyperalgesia (76) and peripheral neuritis (48). Indeed, in most instances the pricking pain threshold was elevated. Yet bedside experience forces recognition of the fact that pin prick or manipulation of the skin caused pain at a lower threshold. When such a patient with nerve injury due to a tumor pressing upon the cervical roots, was carefully studied it was demonstrated that the threshold of burning

pain was actually much lower than normal, whereas that of pricking pain, as previously observed, was considerably higher. It would appear that the nerve itself, because of damage, functioned abnormally. Motivated by the experience of Lehmann (77) that faulty oxygenation of nerves might lower their threshold and even cause spontaneous discharges, burning and pricking pain thresholds were assayed in subjects during periods of nerve asphyxia. Such asphyxia was achieved by binding a blood pressure cuff about the arm, inflating it to 200 mm. of mercury pressure and then during the ensuing forty minutes ascertaining the pricking and burning pain thresholds every few minutes in the manner described above.

Following the onset of asphyxia the pricking pain threshold was slightly lowered for a short time and then started to rise and continued to rise rapidly until in about a half hour pricking pain was no longer discernible. On the other hand, the burning pain threshold was lowered for a much longer time so that at the end of twenty minutes it was thirty-two per cent lower than its control level. It was during this period that "hyperalgesia" existed in the sense that a pin pressed into the skin caused burning pain with its characteristic features. This is the period during which Lewis' and Pochin's "second pain" (75) dominates and the "first pain" disappears (26, 75). It is also comparable to that phase of nerve disease during which true hyperalgesia may exist. This may be expressed for superficial pain in the reaction to the pin as just described, and for deep pain by the tenderness associated with palpation of the deep structures.

In summary, then, it may be said that the skin is endowed with two types of apparatus for perceiving noxious stimuli, one giving rise to the quality of pricking which reaches the sensorium rapidly, the other giving rise to the quality of burning which reaches the sensorium less rapidly. The latter is similar to the quality of pain experienced on the glans and on some mucous membranes. This second quality would appear to be an intermediate between superficial and deep pain.

Inference Concerning the "Hyperalgesia" of Peripheral Neuropathy. The sensory changes that occur during ischemia of an extremity resemble in many respects those which occur in the peripheral neuropathy of alcoholism with nutritional deficiencies (78), as well as in other varieties of peripheral neuropathies. Sensations subserved by the myelinated fibres are impaired before those subserved by the unmyelinated fibres.

Wortis, Stein and Joliffe (78) suggested that such "paradoxic pain" results because the loss of one type of pain sensation enhances the perception of another and supported this view by citing Gasser (73, 74), i.e., "that the function of the impulses which run on ahead of the others is to adjust the excitability of the synapses in preparation for the arrival of the later impulses". According to the views of Wortis, Stein and Joliffe, if the impulses from myelinated fibres are lacking, there would be no consequent decrease in synaptic excitability and therefore the neural impulses incident upon the synapses would give rise to more intense sensation. Gasser (73, 74) concerned himself not with sensation, but

with a spinal reflex pattern. He did not infer that a rapidly traveling impulse, as a result of noxious stimulation, alters a synaptic junction for subsequent impulses from the noxious stimulation so that the latter give rise to a sensation of a different intensity or quality.

A further criticism of the view of Wortis, Stein and Joliffe (78) concerning the possible origin of the dysesthesia in peripheral neuritis is contained in the observation of Gordon and Whitteridge (79). The latter have recorded by means of the electroencephalograph the time interval between onset of painful stimulation and disturbance of alpha rhythm of the cerebral cortex. In control observations, the disturbance of alpha rhythm associated with pain from the finger began 0.26 second after onset of stimulation, and following prolonged asphyxia of the finger, the interval between onset of painful stimulation and disturbance of alpha rhythm was 1.02 seconds. Whereas patients with cutaneous dysesthesia accompanying peripheral nerve lesions in the upper extremity demonstrated no abnormal delay, the average time interval between onset of stimulation and disturbance in alpha rhythm being 0.21 second indicating that the fast conducting pain fibres are still active in these patients. Such data fail to support the concept that dysesthesia results from destruction of fibres conducting fast impulses.

The data from the experiments of Bigelow, Harrison, Goodell and Wolff (48) suggest a simpler conception of the "paradoxic pain" of peripheral neuropathy; as the expression of a defect in peripheral nerve, the threshold for "burning" pain subserved by unmyelinated fibres is so depressed that ordinarily innocuous stimuli are perceived as painful, whereas at the same time the threshold for "pricking" pain, subserved by myelinated fibres, is elevated. Under these circumstances, contact with the bed clothes produces a painful burning, and yet pin prick perception is impaired.

Deep Pain—Visceral and Somatic. Since the head (80) and abdomen (81) are responsible for a large share of human discomfort, and, as well, afford examples of two entirely different mechanisms of visceral pain, consideration will be focussed on these structures.

Pain from the head does not emanate directly from the parenchyma of the brain but from its coverings and from its supporting and vascular structures (18). The brain parenchyma and the linings of the ventricles of the brain are insensitive. The analysis of blood vessel pain reveals that a good portion of the pain from the head has its origin in or about these nutrient structures. For analysis of such blood vessel pain from the head, the cranium may be divided into three layers, the most superficial including vessels on the outside of the head, the middle including the dural vessels, and the deepest layer including the pial and cerebral vessels. Postponing consideration of the pain sensitivity of venous structures, that of the deepest layer of arteries may be analyzed first.

It has been possible to demonstrate in conscious and co-operative individuals that the arteries of the brain are pain sensitive in certain regions, notably at the base (18).

Secondly, it has been shown that histamine which widely dilates cerebral

arteries is capable of producing pain (18, 82, 83, 84). The resultant headache has a pulsatile, throbbing quality. It has been demonstrated that it results from dilatation of the cerebral arteries.

It is well known that such indirect effects as muscle ischemia resulting from vasoconstriction may cause pain (26). Also, it is possible that pain may result indirectly from nerve or root ischemia (85). However, with the possible exception of cold pain (86) there is no evidence to show that arterial contraction causes pain directly. Epinephrine, when locally applied to the surface of the pain sensitive middle meningeal artery caused its walls to contract so vigorously that the end result was a vessel which had but a fraction of its former diameter, yet no pain was elicited by such spasm (18). That constriction of cerebral arteries sufficient to impair brain function does not elicit pain was also indicated by patients with scotomata as forerunners of migraine headache attacks (87). Marked visual field defects in such patients preceded the onset of headache whereas overcoming such constriction by suitable vasodilators, when carried to excess, resulted in pain. On the other hand, migraine headaches due to cranial arterial distention were reduced in intensity or abolished by vigorous vasoconstriction as, for example, with ergotamine tartrate or epinephrine (88, 89).

Since blood vessels entering and leaving the brain act in part as supporting or anchoring structures, they may become the origin of pain through traction upon them or displacement. Many such supporting structures have pain fibres on their surfaces. The method of analysis of such deep pain due to traction was as follows (90): A needle in the subarachnoid space was attached to a water manometer. It was then possible to vary at will the pressure within the subarachnoid space, and to estimate the intracranial or vertex pressure. Under these circumstances traction on anchoring venous structures could be induced. Often, pain could be evoked in a subject with increased intracranial pressure by lowering the pressure to normal. Inversely, a headache associated with low intracranial pressure could be reduced in intensity by raising the pressure to normal. Thirdly, headache failed to be induced in normal subjects by raising the pressure eight or ten times above the normal level (91).

Further, it was observed after spinal drainage and tilting of the subject toward the upright position, when the disparity in negative pressure between the intracranial venous pressure and that in the subarachnoid space became sufficiently great, that the cerebral veins dilated and pain was experienced. Increasing the cerebral venous pressure by jugular compression increased the intensity of this pain. Hence, it is necessary to infer that changes in cerebrospinal fluid pressure, per se, are not responsible for headache. The latter has its origin in traction on pain-sensitive vascular structures (90, 91).

The Nonspecific Quality of Deep Pain. Attempts were made to differentiate in the same person the quality of pain: *a*, in headache resulting from intravenous histamine, causing noxious impulses from cerebral artery distention; *b*, in migraine headache emanating chiefly from distended branches of the external carotid artery; *c*, in headache due to venous traction arising after spinal fluid drainage, and *d*, in headache resulting from injection into the frontalis muscle

of a hypertonic salt solution. Although these headaches varied as regards site, pulsatile or nonpulsatile nature and other temporal aspects, and as to whether or not they could be modified in intensity by such factors as position, movement and chemical agents, the quality of the pain was reported to be the same in all. It was of a deep, aching, diffuse nature, quick to arouse reaction in muscles and in glands, such as lacrimation and injection of the eyes, and was associated with feelings of nausea. It was indistinguishable in quality from that which results from painful stimulation of the teeth, the periosteum, or the muscles of the eye (92).

True Visceral Pain. The evidence concerning true visceral pain has been repeatedly challenged since it is difficult or impossible to stimulate one of the visceral organs without at the same time stimulating supporting structures. Mackenzie (93) emphasized that no conclusion concerning pain sensitivity of an organ should be drawn if stimulating pressure is exerted through it against structures which may in themselves be sensitive. Lewis asks whether it is possible that tenderness in patients with gastric ulcer sometimes emanates from the attachments of the stomach to the posterior abdominal wall. Resolution of this question, says he, is important to the problem of visceral pain, for if one admits that true visceral tenderness occurs, then one simultaneously agrees that pain can result from stimulation of a viscus.

The stomach offers an excellent opportunity to resolve the question as to whether or not there be true visceral pain. Wolf and Wolff studied a man who had a large gastric stoma, surgically produced because of esophageal occlusion when the subject was a child. It was observed in this subject that when 50 or 95 per cent alcohol was introduced into the cardiac end of the esophagus it produced a sensation of "heartburn" (81). These observations confirm those made earlier by Jones and others (24, 94) in individuals with intact gastrointestinal tracts.

When the healthy mucosa of the fundus of the stomach was explored through the stoma described above, and the wall of the mucosa was pressed between the blades of a forceps, no pain resulted. Furthermore, faradic stimulation intense enough to cause pain on the tongue, when applied to the fundus mucosa produced no pain. Fifty and ninety-five per cent alcohol, 1.0 N hydrochloric acid, 0.1 N sodium hydroxide or 1:30 suspension of mustard when applied to the mucosa produced no pain.

However, when the gastric mucosa was inflamed, congested and edematous for whatever reason, all of these procedures evoked a pain of considerable intensity.

Whereas mucus is a highly protective substance, the absence of mucus does not explain the change in pain threshold. When the mucus was aspirated away from the normal mucosa, powdered mustard placed on the now dry and unprotected mucosa did not immediately produce pain. It was not until some time had elapsed and the mucosa had become red and edematous that painful sensations could be elicited by pinching or by faradic current. Also, various noxious chemical agents such as 1.0 N hydrochloric acid and 50 and 95 per cent alcohol now became capable of evoking painful sensations.

Thus, although the number of pain fibres and endings may be relatively small, there can no longer be any doubt about true visceral pain originating in noxious stimulation of the inflamed gastric mucosa.

The question whether pain stems directly from the muscularis or rather from the mesentery and visceral peritoneum is less readily answered. Lewis (26), from his summary of available evidence, inferred that contraction of the muscle of the gut does not give rise to painful sensations. It is his opinion that experiments aimed to resolve this question lead to the conclusion that mesenteric attachments are the sources of pain. It is generally agreed that the parietal peritoneum when suitably stimulated gives rise to sensations of pain.

Recent evidence (81) on the aforementioned subject has come from two series of experiments. 1. Ordinarily when the stomach contracts with a force sufficient to support 35 mm. of mercury pain is elicited. When the gastric mucosa is inflamed it required a force sufficient to support only 20 mm. of mercury to produce a similar pain. It may not be inferred from such an experiment that the pain is elicited exclusively from the mucosa. It is doubtful, moreover, whether it represents mesenteric traction since there is less traction as the result of a force supporting 20 mm. of mercury than 35 mm. It seems likely, therefore, that some of the pain emanates from the structures deeper than the mucosa, namely, from the muscularis, the serosa or the visceral peritoneum. 2. When the stomach was stretched by glass rods it was found that when the contractile state of the stomach was average the pressure necessary to produce pain was 100 grams per sq. cm. However, when the stomach wall was strongly contracted pressure of 50 grams per sq. cm. or half that originally applied was found sufficient to produce pain. On the other hand when the stomach was relatively relaxed 150 grams per sq. cm. of pressure was necessary to induce pain. Again, these observations suggest that the muscularis or serosa may contribue to pain experienced from the stomach.

To supplement these observations on the exposed gastric mucosa, indirect but significant evidence comes from the observations of Palmer (95) who has shown that pain from peptic ulcer occurs only when the gastric content is acid. It is reduced in intensity or eliminated by emptying the stomach as by emesis or aspiration, and by neutralizing the acid content with food or alkali. Also, in the patient with peptic ulcer, who is temporarily free of pain, injections of dilute hydrochloric acid (0.25, 0.5 and 1.0 per cent) will induce pain. Pain is not induced by the injection of such hydrochloric acid into the normal stomach or into the stomachs with healed peptic ulcers.

Also, evidence from x-ray examination, according to Palmer (95), demonstrates that the pain of ulcer is not dependent on hyperperistalsis, sustained contraction of the musculature, pylorospasm or distention of the antrum. However, all of these in conjunction with hyperacidity and inflammation augment the pain of ulcer. Thus, peristaltic contractions may induce pain in a patient with ulcer when there is acid present but if the acid is neutralized, peristalsis no longer induces pain. As in the skin and in the mucosa of the bladder and

of the nose, (96, 99) inflammation of the peptic mucosa also lowers its pain threshold (81).

An Analysis of Deep Pain. A conception that aims to organize the data on deep pain must include an appreciation of the following characteristics: associated muscle contraction either transient or sustained, coupled with smooth muscle and glandular effects; deep tenderness; faulty localization of pain; spread of pain; and surface hyperalgesia and hyperesthesia.

a. Rigidity and Tenderness. It is a common observation that deep pain is followed by local and sometimes distant contraction of skeletal muscle (26, 96). The head is a suitable place for the demonstration of such muscle rigidity and tenderness in association with visceral pain (97, 98).

When the pain sensitive structures about the brain or other deep structures of the head are stimulated, the muscles of the head contract. If this stimulus is short-lived, muscle effects are also brief. If, on the other hand, the painful stimulus from deep structures persists, long-lasting contractions of the muscles of the head and neck and even of the jaws and face take place. Such sustained contraction may, in itself, induce pain, and, as well, tenderness of the muscles involved (for example, in certain patients with brain tumor). Recent experimental evidence permits a more precise statement as to the degree and nature of such pain of muscle origin.

Deep noxious impulses with pain from the head cause contractions especially of the frontal, masseter and temporal muscles, greatly accentuate winking and ultimately cause contractions of occipital and cervical muscles. There may also be vasoconstriction or ischemia in the area of muscle contraction. Lacrimation, nasal congestion, edema of the lids, injection of the conjunctiva, sweating and photophobia, nausea and vomiting are common.

Similar contraction effects followed the introduction of a foreign body into the conjunctival sac or on the cornea. They followed diplopia experimentally induced through the use of lenses (19). Long-standing paranasal disease was also associated with sustained painful contractions of the head and neck muscles (97). Muscle pain and tenderness resulted from painful distention and faradic stimulation of the ureter (96). Occasionally after a sustained and painful distention of the ureter, muscle contraction effects increased to such a degree as to become the major factor in the individual's discomfort for as many as twenty-hours. It is thus apparent that noxious impulses, whether they emanate from somatic or visceral structures give rise to the same type of muscle contraction effects, a conclusion originally formulated by Lewis from his experience with abdominal structures and pain (26).

Lewis and Kellgren (28) have demonstrated the non-specificity of such muscle reactions and also how the muscles themselves may become sources of noxious impulses resulting in pain. Furthermore, it has been shown that the pain due to such sustained contraction of skeletal muscle is ended by disrupting the mechanism of noxious stimulation in the muscle by procaine infiltration (97, 98).

Though muscle tenderness is often a sequel of the prolonged muscle contraction perhaps coupled with ischemia, (26, 75) not all subcutaneous or muscle tenderness

has such an explanation. The experiments of Robertson, Goodell and Wolff (41) demonstrate another mechanism of such deep tenderness. After inducing headache, deep tenderness, surface hyperalgesia and hyperesthesia in the temporal region of the head by prolonged stimulation of a tooth, procaine was infiltrated intracutaneously at the site of most intense temporal headache, tenderness and surface hyperalgesia. Within a few seconds after the injection there was analgesia of the skin and the subject estimated that a portion of the headache was eliminated. The temporal muscle beneath the analgesic skin remained tender to palpation. When procaine was injected into the belly of the temporal muscle as well as intracutaneously, pain still persisted although local tenderness was eliminated. However, when the tissue about the noxiously stimulated tooth was infiltrated with procaine, there was complete elimination of all pain.

It is therefore suggested that tenderness sometimes results from central spread of excitatory effects of sustained noxious stimulation. A similar mechanism is invoked to explain surface hyperalgesia and hyperesthesia and is considered further in a following section.

b. *Localization and Spread.* The localization of deep pain is conveniently studied by stimulation of the nasal and paranasal structures (99). Thus, when the mucosa about the ostium of the maxillary sinus was stimulated either mechanically or by faradic current there resulted at first a localizable intranasal painful sensation. When the noxious stimulation was continued, the subject experienced pain which spread over the homolateral portion of the nose and cheek, along the zygoma, into the temporal region, and into the upper teeth. In short, the pain spread at first from the site of stimulation to predictable sites in the same segment. When noxious stimulation of the turbinates was continued for ten minutes the area of pain spread over most of the area of distribution of the second division of the fifth cranial nerve and ultimately spread to involve adjacent portions supplied by the third and first divisions. This new, more widespread area of pain had a deep aching quality, which became so intense as to minimize the pain from the nose.

The deep aching sensation was associated with blushing of the skin over the cheek, injection of the conjunctiva, lacrimation and photophobia of the homolateral eye. Hence, although deep pain may be initially localized and correctly identified as to site of stimulation, subsequently it may become falsely localized because it is experienced far from the site of stimulation. Also, noxious stimulation of widely separated structures of the nasal and paranasal region caused an identical distribution of pain in remote portions of the face or head. Depending upon the intensity and duration of the painful stimulation the pain may remain localized within the same division or segment, or it may spread so that it seems to emanate from peripheral structures innervated by adjacent divisions or segments.

The localization and spread of pain resulting from stimulation of nasal and paranasal structures are typical of deep pain. Thus, pain from the stomach mucosa, although it is experienced initially as from the stomach itself, subsequently seems to emanate from other parts of the same and adjacent segments.

The latter sensations may dominate so that the initial circumscribed pain is less apparent.

McLellan and Goodell (96) have shown that brief low intensity electrical stimulation of the ureter from within elicits prompt and severe pain experienced along the medial border of the rectus abdominis muscle. Spread from an initially limited portion of a segment to other portions of the segment and then ultimately into adjacent segments was also experienced when the ureter was stimulated at various levels above the bladder orifice. When such stimulation was prolonged or intense, the pain spread so as to include other parts of the segment, beginning on the anterior abdominal wall and ultimately spreading posteriorly. However, the effects of muscle contraction soon masked the initial pain from noxious stimulation of the ureter to such an extent that the entire flank became tender. The muscle contraction then apparently became the dominant source of noxious impulses resulting in pain.

Similar observations were made by Travell, Berry and Bigelow (100, 101) who mapped the distribution of referred pain from noxious stimulation of various muscles. They observed that the referred pain was experienced either within the reference area attributed to one segment, within the different portions of one segment or in fragments of several segments without including any one segment entirely.

With coronary artery occlusion or angina of effort the pain at first seems to emanate from one or more of the anterior portions of the first four or five thoracic segments and spreads headward so as to seem to arise within the structures supplied by the lower cervical, then the upper cervical segments. Ultimately it may seem to emanate from the lower jaw and teeth. The spread of pain so as to include the lower teeth involves the cervical dorsal horn and the descending nucleus of the trigeminal nerve which are contiguous in the upper cervical cord.

To further the study of mechanisms involved in the distribution of referred pain, the commonly experienced "ice cream" headache was analyzod in a subject in whom it was possible to examine separate parts of the esophagus, stomach and mouth (102). It was found that frontal headache developed when the ice was applied to the roof of the mouth but not when a considerable amount of ice was put into the stomach through a stoma, nor when it was held for several minutes in the esophagus. Thus, noxious stimulation of structures supplied by one division of the fifth cranial nerve caused pain to be experienced in structures supplied by another portion of the fifth cranial nerve. Also, but less constantly, pain was experienced in and behind the ear when ice was applied to the region of the fossa of Rosenmuller, and to the posterior wall of the pharynx. These areas are all supplied by the ninth and tenth cranial nerves with possibly a few contributions from the fifth and seventh cranial nerves.

The digits of the hand in certain persons are particularly suitable for the analysis of the spread of pain from a site of noxious stimulation because the pain can be recognized in readily separable zones.

The "cold pain" induced by immersing one digit for 10 minutes in water at

0°C. spreads from the immersed finger to the adjacent border of the neighboring fingers on either side, and may subsequently include the whole of a neighboring finger and a part of a finger beyond.

This pattern of spread having been well defined in 21 experiments in 6 subjects, the fifth finger was anesthetized by a digital block with 2 per cent procaine. The infiltration of the paired dorsal and volar digital nerves was extended to encircle the base of the digit and included dermis and periosteum. The resultant sensory loss in this finger was complete, with absence of superficial and deep pain and of touch, temperature, position sense and vibration sense. When the fourth finger was immersed in water, the spread of pain to the fifth and third fingers was essentially of the same pattern and of the same intensity as in control experiments in which the sensation of the digit was intact. Repetition of this experiment with procaine block of the third finger and immersion of the second in cold water also revealed that there was no interference with the spread of pain into the anesthetic digit. It is inferred that *a*, the spread of pain is a central rather than a peripheral effect; and *b*, that the spread is not dependent upon afferent impulses from the tissues into which spread occurs (103).

Such segmental spread of the effects of noxious stimulation associated with deep pain would appear to be based chiefly on anatomical arrangements within the spinal cord or brain stem. Thus, in one of two similar patients studied by Ray and Wolff (104) severe pain in the right buttock and lower leg, caused by metastatic tumor of the right hip joint and nerve plexus, was eliminated by section of the ventrolateral portion of the cord at the first thoracic segment, on the left side only. After operation, severe deep pain induced by compressing muscles on the left (normal) lower limb or abdomen was felt on the left side only. Similar noxious stimulation of the right hip joint or right gastrocnemius muscle induced a less intense deep pain but also on the left side; the pain was diffuse, poorly localized and widespread, yet mainly in the tissues supplied by the adjacent segments, and it outlasted the period of stimulation by about half a second. Noxious stimulation of low intensity, or repeated pin pricking on the right side gave no sensation of pain on the left. It is postulated from these observations that excitation travels along two possible routes from the primary sensory neurons and their collaterals: 1, the well-known pathway, by neurons in the cord segments of their entry and in adjacent segments that cross to the opposite side and ascend the cord in the spinothalamic tract; and 2, the pathway demonstrated by this patient, that of internuncial neurons that cross in the posterior commissure to connect with neurons on the opposite side of the cord. The latter cross to ascend the cord in the spinothalamic tract on the same side as the entering noxious impulses. Pain is thus experienced on the side opposite to the source of noxious impulses. It is conceivable that the additional synapses necessary for the passage of impulses from the side of entry to the other side of the cord introduces resistance that makes necessary a high intensity of noxious stimulation for such spread of pain to the opposite side of the body.

Included in the evidence of the effects of spread from noxious stimulation on sensory and motor phenomena may be cited the observations of Gellhorn (105,

106) who has shown that spinal reflex movements, as well as those resulting from stimulation of the motor areas of the cerebral cortex are altered during periods of noxious stimulation of either deep or cutaneous tissues. During such noxious stimulation the effect of stimulation of the motor cortex is either intensified or qualitatively altered so that different movements may be substituted temporarily for those obtained under control conditions. These afferent impulses are not limited in their effects to the same side but may also involve the reactivity of the opposite side. Gellhorn suggested that these phenomena result from the cortical spread of excitatory effects, but the evidence for this inference is inconclusive. The effects described in his experiments could be explained by assuming spread at the spinal cord level.

Headache occurs in association with distended bladder in some paraplegic patients with verified cord transection, usually during the first few months following injury. Such headaches are described as throbbing and they may be diffuse, bifrontal, or bitemporal. They are generally accompanied by diffuse sweating and piloerection *above* the level of cord transection as well as below. In most instances the headache is relieved almost immediately after the bladder has been drained. It is likely that such headaches have a vascular mechanism, but the anatomic pathways involved have not been ascertained.

Also, it is difficult to understand on the basis of neural connections the instances presented by Henry Head (107), which involve headache in association with chest disease. Persons with pain emanating from one portion of the chest were reported to have had headaches in one specific portion of the head, whereas those with disease and pain in another portion of the chest had headache in an entirely different site. It is apparent from the case notes that these patients had septicemia or bacteremia and the headache could have resulted from humoral agents acting on cerebral vessels in much the same manner as has been described for experimentally induced histamine or typhoid fever headaches (82, 83, 84, 89). Assuming, however, that the humoral interpretation is valid, the predictable and specific site of the headache is still unexplainable. Others have been unable to confirm Head's observations.

The site selected in the faulty localization of pain appears to depend first of all on the familiarity of the subject with painful experience from a region. Thus, on stimulation of intracranial structures, those near the orbit and those as remotely removed from the orbit as is the superior surface of the tentorium gave rise to pain experienced in, over or behind the eye. If pain be "referred to" or experienced in the part of the segment most accustomed to noxious stimulation it must emanate from the region about the eye since here dust, foreign bodies, drying, inflammation and muscle contraction are common causes of noxious stimulation. It is likely that the cerebral cortex is involved in those aspects of pain that include previous experience or conditioning.

c. Surface Hyperalgesia and Hyperesthesia. On certain portions of the skin, as an accompaniment of deep pain, pin prick may be experienced as sharper and of longer duration, and tactile and thermal sensations seem more intense than those induced by the same stimuli on other dermatomes. Such skin areas

are innervated by the same or adjacent neural segments which supply the deep tissues from which the noxious and painful impulses are originating.

This interesting yet inconstant accompaniment of deep pain has caused confusion and controversy. Though described many years ago (108), the phenomenon was given fresh interest by Weiss and Davis (109). These authors suggested that deep pain could be reduced or eliminated by anesthetization of the associated hyperalgesic areas of the skin. Morely (110) reported that hyperalgesia of the skin over the shoulders from diaphragmatic irritation could be eliminated or reduced by anesthetization of the hyperalgesic area, though Woollard, Roberts and Carmichael (111) demonstrated that the more intense pain resulting from direct phrenic nerve stimulation and experienced in the shoulder tip was in no way influenced by surface anesthetization of this area.

Lewis (26) was not able to modify the pain of angina pectoris by anesthetization of the precordial skin nor was it possible for McLellan and Goodell (96) to alter the pain experience associated with ureteral distention by anesthetization of the hyperalgesic skin area of the abdominal wall. Neither could pain resulting from duodenal distention be modified by anesthetization of the hyperalgesic area in the skin of the abdominal wall (112). Anesthetization of the skin over the calvarium (not hyperalgesic) in no way influenced the intensity of the headache resulting from intravenous histamine injection (113). Also, as mentioned above, (103) anesthetization of the digits in the zone of reference did not modify the pain (see p. 187).

These contradictions are more apparent than real. It was shown by Robertson, Goodell and Wolff (41) that the effect on deep pain of procaine infiltration is related to the occurrence and amount of surface and deep hyperalgesia. Thus, when a tooth was noxiously stimulated causing headache and superficial and deep hyperalgesia of the temporal region of the head, infiltration of procaine into the hyperalgesic skin and underlying soft tissues reduced the amount of discomfort and produced analgesia although it did not eliminate the headache. When the hyperalgesia was more marked the effect of procaine injection was more dramatic. Headache was eliminated, however, by infiltration of procaine into the tissue about the tooth (41). It is, therefore, evident that when pain results from the persistence of primary visceral or other deep noxious stimulation and is associated with hyperalgesia, its intensity may be modified by superficial and deep procaine infiltration in the hyperalgesic zones. However, it is not eliminated until the afferent impulses from the primary sources end spontaneously or are blocked.

The explanation of surface hyperalgesia presented by Mackenzie (93), Ross (114) and Sturge (108) included autonomic afferents and "hyperirritable" foci within the cord. Lewis (26) has postulated that a special nervous apparatus called the noci-fensor system is involved in the phenomenon of hyperalgesia, but anatomic evidence for the existence of such an apparatus is lacking.

It is extremely doubtful that there is liberated in the skin of the hyperesthetic and hyperalgesic areas a chemical agent which lowers the pain threshold (115). Repeated observation of patients with surface hyperesthesia and hyperalgesia

associated with deep pain reveals no significant lowering of the threshold of either pricking or burning pain. On the other hand, minimal pricking and burning pain when they are perceived are experienced as more intense, and seem to be of longer duration. In other words, there is apparently an intensification of the pain perceived at the usual rather than at a lowered threshold (48, 76, 115).

For example, a man with hyperalgesia of the right side of the face, secondary to noxious impulses from deep tissues, was found to have the same pain threshold on the two sides of his face, which was within normal limits. However, a stimulus inducing a sensation which was reported as one plus pain on the control cheek and forehead, induced a sensation which was reported as a three plus pain on the hyperalgesic side of the face. A stimulus inducing a sensation reported as a four plus pain on the control side of the face induced a sensation reported as a six plus pain on the hyperalgesic side. Also, a stimulus of 0.295 gram cal./sec. on the hyperalgesic side induced a sensation which was reported as equivalent in intensity to one produced by a stimulus of 0.335 gram cal./sec. on the control side (115).

If there be liberated into the skin substances capable of lowering the threshold, then the effects of such substances should persist in the skin for at least a short time after the stimulus liberating the substance has ended. To explore this thesis the following experiment was done: An area of hyperalgesia (intensification of prick with pin and touch with cotton wisps) was produced on the skin of the left cheek over the outer edge of the zygoma by placing under the left middle and inferior turbinates an irritating "adrenalin" (Parke, Davis & Co.) tampon (1:1000 solution). The pain threshold for both pricking and burning pain in the hyperalgesic area, as ascertained by the thermal radiation technique, was not lowered. Moreover, when the pain in the face had been obliterated by placing procaine tampons over the irritated surface of the nasal mucosa, almost immediately thereafter the "hyperalgesia" and "hyperesthesia" of the skin was eliminated. In other words, the "hyperalgesia" and "hyperesthesia" did not outlast the period of painful stimulation. These observations make it seem extremely unlikely that a chemical agent was liberated into the skin, which either produced true surface hyperalgesia, or was responsible for the intensification noted (59, 116).

It appears more likely that the intensification of sensation results from a change within the nervous system. It is suggested that the central excitory effects due to the pre-existing barrage of noxious impulses in the segmental or suprasegmental neural apparatus alter the situation so as to make impulses originating at the usual threshold in the skin seem more than usually intense and persistent.

A summation effect analogous to the phenomenon of surface hyperalgesia is encountered in the eye. It is common experience that persons with a cinder in the eye will, on looking into a light, experience the light as momentarily brighter and the cinder as more painful. It is as though a mutual reinforcement of visual sensation and pain occurs (59).

The following experiment may make this more evident (59). There was injected into the muscles of the forehead a small amount of hypertonic saline (0.3 cc. of 6 per cent solution). This caused intense pain. The rate of winking was frequently measured before and after the injection. It was noted that the rate multiplied seven to eight times when the eye was exposed to a light of standard intensity, and dropped to zero when the subject was put in a dark room. Also, immediately after exposure the light seemed brighter than during the control period and the pain in the head due to the salt injection was increased in intensity.

A possible explanation of this phenomenon is as follows: At the segmental level visual impulses entering the colliculus from the retina, exert an excitatory influence on the facial nerve nucleus, causing the increase in winking frequency The spread of the pain from the site of noxious impulses in the muscles of the forehead to other parts of the head results from spread of excitation to involve most of the trigeminal nerve nucleus so that the precise peripheral origin of the pain is no longer certain. Noxious impulses from the head after entering and spreading throughout the trigeminal nucleus exert an additional excitatory influence on the facial nerve motor nucleus so as further to increase the winking frequency (117). The spread of excitation both caudad and cephalad causes the facial, cervical, masseter and temporal muscles to contract. Additional evidence that the motor effects of this experiment are segmental (midbrain or hindbrain) in origin was afforded by some patients with Argyll-Robertson pupil (indicating mid-brain disease). Such persons had faulty or absent wink responses when a strong light was projected onto the retina. Yet they also experienced mutual intensification of light and pain.

The mutual intensification of pain and visual sensations must involve association areas in the cerebral cortex. Nervous impulses from stimulation of the retina by light, entirely separate from those which go to the colliculus, enter the external geniculate body directly. There, through a synapse, secondary disturbances are conveyed to the cerebral cortex. Assuming that current knowledge of neuroanatomy is correct, it is apparent that at neither the brain stem nor thalamic level is it possible for neural activities involved in vision to be influenced or modified by those for pain. Such phenomena can take place at the cortical level only (118). Thus, although reinforcement of the effects of noxious stimuli by non-noxious stimuli does occur within the cerebral cortex, it is also likely that such effects do occur within the cord or brain stem.

To epitomize this analysis it may be said that there are three categories of deep pain:

1. True visceral and deep somatic pain. Such pain is felt at the site of primary stimulation and may or may not be associated with referred pain. It is eliminated by infiltration of procaine into the site of noxious stimulation or by blocking its afferent nerves, but it is not altered by infiltration of procaine into other structures supplied by the same or adjacent neural segments.

2. Referred pain. Such pain may occur in addition to or in the absence of the true visceral and deep somatic pain described above. It is experienced at

a site other than that of stimulation but in tissues supplied by the same or adjacent neural segments. It may occur either with or without associated hyperalgesia and hyperesthesia.

a. Without superficial and/or deep hyperalgesia. In this case pain depends only on the central effects of the spread of excitation of the original noxious impulses to the same and adjacent segments of the cord whence they are relayed to higher centres for perception and interpretation. Injection of procaine into superficial or deep regions of referred pain does not reduce the intensity of pain due to this mechanism.

b. With superficial and/or deep hyperalgesia. Referred pain may be accentuated in intensity by virtue of the effects of ordinarily non-noxious stimuli from zones of reference. Impulses from such sources, normally inadequate to produce pain, may do so upon reaching the cord in a segment involved in central spread of excitation. Procaine injected into superficial or deep hyperalgesic structures will abolish this element of the referred pain phenomenon, resulting in more or less reduction of the subject's discomfort, depending on the amount of hyperalgesia.

3. Pains due to secondary skeletal muscular contractions which provide a fresh source of noxious impulses. Pain may result from secondary effects of the central spread of excitation on the effector structures, including painful contractions of skeletal muscles. Such disturbances may be widespread and the pains may be experienced in situations remote from the original source of noxious stimuli. Local infiltration of the contracted muscles with procaine abolishes this type of pain by disrupting its peripheral mechanism.

Thus in addition to the above described (2*a*) hyperalgesia with unaltered cutaneous pain threshold, there may occur secondarily a deep hyperalgesia with lowered pain threshold in the deep tissues.

The Phenomenon of "Central Pain". The phenomenon of central pain has been but little investigated experimentally. Head and Holmes (119) explained this dysesthesia by postulating the failure of cortical inhibition of excitation within the thalamus, due to lesions that destroyed cortical thalamic connections. Kendall (120) has suggested that the dysesthesia (the phenomenon of "central pain") associated with lesions of the spinothalamic tract and of the thalamus can be interpreted as the result of differential interruption of pathways for afferent impulses inducing pain. He postulates that it is the fast conducting fibres in the spinal cord and in the thalamus which are interrupted by a lesion which leaves intact the slow conducting fibres, the central effects of which, it is further postulated, are normally inhibited by the effects of impulses traversing the fast fibres. The fact that pricking pain is experienced by those with central pain, although its threshold is elevated (48), fails to support such a view.

The Dual Aspects of Pain. The distinction between perception and reaction is apparent and is easily appreciated in the case of heat, light, touch, cold and olfactory perception where the responses are not stereotyped. But when, as in the case of pain, there is an inborn as well as an acquired stereotypy the contrast between perception and reaction is blurred. The reaction pattern of the organ-

ism to noxious stimuli involving pain has many components. It includes "feeling" or emotional reactions, smooth muscle, gland and skeletal muscle effects, and assumes the form of withdrawal, flight or fight.

Instances of the dissociation of pain perception from the pattern of reaction to pain are seen in the indifference to injury sustained during the excitement of games or combat; the apathy, or "quietism" that accompanies tissue damage during certain religious and mystical practices; the indifference to tissue damage during sexual excitement, the indifference to pain often witnessed during parturition in women who are confident in their physician and desirous of bearing a child (14).

Dissociation between pain perception and pain reaction is evident not only after drug administration or during the action of strong beliefs and convictions, but also after cerebral damage. Examples of such dissociation are offered by the experiences of VanWagenen (121) and Watts (122) and of Walker (123). Thus, Walker's patient, a fifty-six year old woman with intractable pain associated with an amputation stump and "phantom limb", had bilateral frontal lobotomy performed for this complaint. In the months following this procedure she was confused but looked content and complained little. When questioned she usually said the arm "pained" but did not volunteer the information. When she was questioned further about the pain she stated that the pain was present but that she gave it less attention and it did not concern her. (See also de Gutiérrez-Mahoney, 32a.)

The dissociation of perception from reaction may be of varying degrees. There may be more or less denial of emotional reactions with repression of vocalization or flight, yet with many visceral reactions sometimes ending in syncope. Indeed, some of these reactions if sustained may damage the organism. For example, G. Wolf (124) has demonstrated that noxious stimuli associated with pain may induce a decrease in kidney blood flow and urine output. Such ischemia is of little importance to the healthy kidney, but in a severely damaged organ, a transient decrease in kidney circulation has resulted in irreversible changes and death (125). Also, Gold, Kwit and Modell (126) have shown that noxious stimuli associated with pain may alter heart function as indicated by the electrocardiogram. Again, such changes, although of little importance to the healthy heart may be of dire significance to one in already faulty function.

Of the various bodily changes available for the analysis and study of such reactions, a sweating and vasomotor phenomenon as represented by changes in the electrical skin resistance has been found useful (17). The point of change in skin resistance resulting from palmar skin changes when the forehead of the subject was exposed to heat radiation, was called the alarm reaction threshold.

Although the pain threshold was constant, the "alarm" reaction threshold varied widely from subject to subject and in the same subject from time to time. It was usually below the pain threshold, although sometimes above it. This reaction is initiated by painful and nonpainful stimuli, when the latter have become symbols of painful or dangerous experience. The reaction or "alarm" threshold differs from the pain threshold not only in its variability but in the

fact that there is spatial summation. Moreover, an analgesic such as alcohol, which has a uniform and predictable effect on the pain threshold, affects the "alarm" reaction in a variable way in the same and different individuals, raising it in amounts from 50 to 800 per cent. The threshold for reaction has been measured by other methods. Chapman et al. (42) using radiant heat to the forehead as a painful stimulus, evaluated a winking response characterized by narrowing of the eyelids as an evidence of reaction to pain. They found that "neurotic" patients in general displayed a lower threshold for reaction than did "control" individuals (127). The older observations of Libman (44) and of Hollander (45) support this view. This reaction component, then, of the pain experience includes the "quale" or feeling state and represents the individual's response to a given situation based on his own past experience and inborn peculiarities. Impressions concerning the rôle of the internuncial neurons in the mechanism of pain in phantom limb and causalgia, as well as their relevance to other sustained states of abnormal reactivity have been described by Livingston (128).

In concluding these considerations, concepts concerning pain may be touched upon again. Until the end of the nineteenth century pain was considered to be exclusively a feeling state. Later, with the discovery of special anatomic equipment and mechanisms, interest was focussed on the perceptual aspects of pain. It then became clear that pain is a specific sensation, and yet, because of its intimate linkage with strong feelings and other reaction patterns, the latter may be dominant in the experience. Reactions to pain may be modified by conditioning experiences and strong beliefs. The common analgesics in addition to effectively raising the pain threshold, also have a major function in changing reactions, attitudes and feelings. Thus, recent evidence supports the old view that the "quale" or feeling state is, to the one who suffers, perhaps the most relevant aspect of pain. Yet it supports as well the conclusion that pain is a specific sensation with its own structural and functional properties. It becomes apparent that these two concepts do not oppose each other: both represent attempts to formulate distinct but fundamental aspects of the pain experience.

REFERENCES

(1) DALLENBACH, K. M. Pain: History and present status. Am. J. Psychol. 52: 331, 1939.

(2) DANA, C. L. Paper on Syringomyelia: Central glioma of the spinal cord, with spontaneous central hemorrhage. New York Neurological Society Reports of December 5, 1893 meeting. J. Nerv. and Ment. Dis. 21: 67, 1894.

(3) BLIX, M. Experimentelle Beiträge zur Lösung der Frage über die specifische Energei der Hautnerven. Ztschr. f. Biol. 20: 141, 1884.

(4) GOLDSCHEIDER, A. Weitere Mitteilungen zur Physiologie der Sinnesnerven der Haut. Pflüger's Arch. 168: 38, 1917.

(5) VON FREY, M. Die Gefühle und ihr Verhaltnis zu den Empfindungen. Beit. z. Physiol. des Schmerzsinnes. Berichts über die verhandlung d. Königl. sächs. Gesellshaft die Wissenschaften, Leipzig. Math-phys. Kl. 49: 169, 1897.

(6) WOOLLARD, H. H., G. WEDDELL AND J. A. HARPMAN. Observations on the neurohistological basis of cutaneous pain. J. Anat. 74: 413, 1940.

(7) WEDDELL, G. The pattern of cutaneous innervation in relation to cutaneous sensibility. J. Anat. 75: 346, 1941.
(8) WEDDELL, G. The multiple innervation of sensory spots in the skin. J. Anat. 75: 441, 1941.
(9) TOWER, S. Pain: Definition and properties of the unit for sensory reception. A. Research Nerv. and Ment. Dis. Proc. 23: 16, 1943.
(10) HOLMES, G. Contributions to medical and biological research dedicated to Sir William Osler in honor of his seventieth birthday, July 12, 1919, by his pupils and co-workers. New York, Paul B. Hoeber, 1919, p. 235.
(11) STOOKEY, B. The management of intractable pain by chordotomy. A. Research Nerv. and Ment. Dis. Proc. 23: 416, 1943.
(12) ADRIAN, E. D., McK. CATTELL AND H. HOAGLAND. Sensory discharges in single cutaneous nerve fibres. J. Physiol. 72: 377, 1931.
(13) CATTELL, McK. AND H. HOAGLAND. Response of tactile receptors to intermittent stimulation. J. Physiol. 72: 392, 1931.
(14) WOLFF, H. G., J. D. HARDY AND H. GOODELL. Studies on pain. Measurement of the effect of morphine, codeine and other opiates on the pain threshold and an analysis of their relation to the pain experience. J. Clin. Investigation 19: 659, 1940.
(15) WOLFF, H. G., J. D. HARDY AND H. GOODELL. Measurement of the effect on the pain threshold of acetylsalicylic acid, acetanilid, acetophenetidin, aminopyrine, ethyl alcohol, trichlorethylene, a barbiturate, quinine, ergotamine tartrate and caffeine: an analysis of their relation to the pain experience. J. Clin. Investigation 20: 63, 1941.
(16) WIKLER, A., H. GOODELL AND H. G. WOLFF. Studies on pain. The effects of analgesic agents on sensations other than pain. J. Pharmacol. and Exper. Therap. 83: 294, 1945.
(17) WOLFF, H. G., J. D. HARDY AND H. GOODELL. Studies on pain: Measurement of the effect of ethyl alcohol on the pain threshold and on the "alarm" reaction. J. Pharmacol. and Exper. Therap. 75: 38, 1942.
(18) RAY, B. AND H. G. WOLFF. Experimental studies on headache. Pain sensitive structures of the head and their significance in headache. Arch. Surgery 41: 813, 1940.
(19) ECKARDT, L. B., J. M. McLEAN AND H. GOODELL. Experimental studies on headache. The genesis of pain from the eye. A. Res. Nerv. and Ment. Dis. Proc. 23: 209, 1943.
(20) KIESOW, F. The problem of the condition of arousal of the pure sensation of cutaneous pain. J. Gen. Psychol. 1: 199, 1928.
(21) CAPPS, J. A. An experimental and clinical study of pain in the pleura, pericardium and peritoneum. MacMillan Co., New York, 1932.
(22) KUNTZ, A. The autonomic nervous system. Lee and Febiger, Philadelphia, 1934.
(23) WHITE, J. C. The autonomic nervous system. MacMillan Co., New York, 1935.
(24) JONES, C. Digestive tract pain. Diagnosis and treatment. MacMillan Co., New York, 1938.
(25) MOORE, R. M. Some experimental observations relating to visceral pain. Surgery 3: 534, 1938.
(26) LEWIS, T. Pain. MacMillan Co., New York, 1942.
(27) ALVAREZ, W. C. Nervous indigestion and pain. Paul B. Hoeber, Inc., New York, 1943.
(28) LEWIS, T. AND J. H. KELLGREN. Observations relating to referred pain, visceromotor reflexes and other associated phenomena. Clin. Science 4: 47, 1939.
(29) CLARK, D., J. HUGHES AND H. S. GASSER. Afferent function in the group of nerve fibres of slowest conduction velocity. Am. J. Physiol. 114: 69, 1935.

(30) GASSER, H. S. Pain-producing impulses in peripheral nerves. A. Res. Nerv. and Ment. Dis. Proc. **23**: 44, 1943.
(31) WALKER, A. E. Central representation of pain. A. Res. Nerv. and Ment. Dis. Proc. **23**: 63, 1943.
(31a) EVANS, J. P. Quoted by WALKER. (See reference 31.)
(32) MICHELSEN, J. J. Subjective disturbances of the sense of pain from lesions of the cerebral cortex. A. Res. Nerv. and Ment. Dis. Proc. **23**: 86, 1943.
(32a) DE GUTIÉRREZ-MAHONEY, C. G. The treatment of painful phantom limb by removal of post-central cortex. J. Neurosurg. **1**: 156, 1944.
(33) KUNKLE, E. C. AND W. P. CHAPMAN. Insensitivity to pain in man. A. Res. Nerv. and Ment. Dis. Proc. **23**: 100, 1943.
(34) FORD, F. R. AND L. WILKINS. Congenital universal insensitiveness to pain. Bull. Johns Hopkins Hospital **62**: 448, 1938.
(35) BEECHER, H. K. Pain in men wounded in battle. Ann. Surg. **123**: 96, 1946.
(36) HARDY, J. D., H. G. WOLFF AND H. GOODELL. Studies on pain. A new method for measuring pain threshold: Observations on spatial summation of pain. J. Clin. Investigation **19**: 649, 1940.
(37) HARDY, J. D., H. G. WOLFF AND H. GOODELL. The pain threshold in man. A. Res. Nerv. and Ment. Dis. Proc. **23**: 1, 1943. Am. J. Psychiat. **99**: 744, 1943.
(38) GOETZL, F. R., D. Y. BURRILL AND A. C. IVY. A critical analysis of algesimetric methods with suggestions for a useful procedure. Quarterly Bulletin, Northwestern Univ. Med. School, Chicago **17**: 280, 1943.
(39) ROTH, G., D. Y. BURRILL AND A. C. IVY. The effects of histamine administered intravenously in increasing concentration on the pain threshold of normal subjects. Federation Proc. **5**: 89, 1946.
(40) KLEITMAN, N. AND A. RAMSAROOP. Body temperature and cutaneous sensitivity to tingling and pain. Federation Proc. **5**: 56, 1946.
(41) ROBERTSON, S., H. GOODELL AND H. G. WOLFF. Studies on headache: the teeth as a source of headache and other pain. Arch. Neurol. and Psychiat., in press.
(42) CHAPMAN, W. P. AND C. M. JONES. Variations in cutaneous and visceral pain sensitivity in normal subjects. J. Clin. Investigation **23**: 81, 1944.
(43) HARRISON, I. AND N. H. BIGELOW. Quantitative studies of visceral pain produced by the contraction of ischemic muscle. A. Res. Nerv. and Ment. Dis. Proc. **23**: 154, 1943.
(44) LIBMAN, E. Observations on sensitiveness to pain. Trans. Assn. Am. Phys. **41**: 305, 1926.
(45) HOLLANDER, E. Clinical gauge for sensitivity to pain. J. Lab. and Clin. Med. **24**: 537, 1939.
(46) BAZETT, H. C. AND B. MCGLONE. Temperature gradients in the tissues in man. Am. J. Physiol. **82**: 415, 1927.
(47) OPPEL, T. W. AND J. D. HARDY. Studies on temperature sensation—The temperature changes responsible for the stimulation of the heat end organs. J. Clin. Investigation **16**: 525, 1937.
(48) BIGELOW, N., I. HARRISON, H. GOODELL AND H. G. WOLFF. Studies on pain: Quantitative measurements of two pain sensations of the skin, with reference to the nature of the "hyperalgesia of peripheral neuritis". J. Clin. Investigation **24**: 503, 1945.
(49) SCHUMACHER, G. A., H. GOODELL, J. D. HARDY AND H. G. WOLFF. Uniformity of the pain threshold in man. Science **92**: 110, 1940.
(50) SCHUMACHER, G. A. The influence of inflammation on the pain threshold of the skin in man. A. Res. Nerv. and Ment. Dis. Proc. **23**: 166, 1943.
(51) WOLFF, H. G. AND H. GOODELL. The relation of attitude and suggestion to the perception of and reaction to pain. A. Res. Nerv. and Ment. Dis. Proc. **23**: 434, 1943.

(52) Gammon, G. D. and I. Starr. Studies on the relief of pain by counterirritation J. Clin. Investigation 20: 13, 1941.
(53) Andrews, H. L. The effect of morphine and prostigmine methylsulfate. J. A. M. A. 120: 525, 1942.
(54) Andrews, H. L. Skin resistance changes and measurements of pain threshold. J. Clin. Investigation 22: 517, 1943.
(55) Andrews, Howard L. The effect of opiates on the pain threshold in post-addicts. J. Clin. Investigation 22: 511, 1943.
(56) Ivy, A. C., F. R. Goetzl, S. C. Harris and D. Y. Burrill. The analgesic effect of intracarotid and intravenous injection of epinephrine in dogs, and of subcutaneous injection in man. Quart. Bull. Northwestern Univ. Med. School, Chicago 18: 298, 1944.
(57) Burrill, D. Y., F. R. Goetzl and A. C. Ivy. The pain threshold raising effects of amphetamine. J. Dental Research 23: 337, 1944.
(58) Ivy, A. C., F. R. Goetzl and D. Y. Burrill. Morphine dextroamphetamine analgesia. War Medicine 6: 67, 1944.
(59) Wolff, H. G. Some observations on pain. Harvey Lectures 39: 39, 1943.
(60) Hardy, J. D., H. Goodell and H. G. Wolff. Studies on pain: Discrimination of differences in intensity of painful stimuli. In press.
(61) Granit, R. and P. Harper. Comparative studies on the peripheral and central retina. 2-Synapatic reactions in the eye. Am. J. Physiol. 95: 211, 1930.
(62) Oppel, T. W. and J. D. Hardy. Studies in temperature sensation. II. The temperature changes responsible for the stimulation of the heat end organs. J. Clin. Investigation 16: 525, 1937.
(63) Hardy, J. D. Additional observations on thermal sensation and discrimination in relation to intensity of stimulus. In press.
(64) Riesz, R. R. Differential intensity sensitivity of the ear for pure tones. Physical Review 31: 867, 1928.
(65) Stevens, S. S. and A. H. Davis. Hearing. John Wiley and Sons, New York, p. 110, 1928.
(66) Best, C. H. and N. B. Taylor. Physiological basis of medical practice. 3rd ed. p. 1628. Williams and Wilkins Co., Baltimore, 1943.
(67) Gad, J. and A. Goldscheider. Ueber die summation von Hautreizen. Ztschr. f. klin. Med. 20: 339, 1892.
(68) Boring, E. G. Cutaneous sensation after nerve division. Quart. J. Exper. Physiol. 10: 1, 1916.
(69) Heinbecker, P. and G. H. Bishop. The mechanism of painful sensations. A. Res. Nerv. and Ment. Dis. Proc. 15: 226, 1935.
(70) Bishop, G. H. Responses to electrical stimulation of single sensory units of skin. J. Neurophysiology 6: 361, 1943.
(71) Bishop, G. H. The peripheral unit for pain. J. Neurophysiology 7: 71, 1944.
(72) Bishop, G. H. The structural identity of the pain spot in human skin. J. Neurophysiology 7: 185, 1944.
(73) Gasser, H. S. Conduction in nerves in relation to fiber types. A. Research Nerv. and Ment. Dis. Proc. 15: 35, 1934.
(74) Gasser, H. S. The control of excitation in the nervous system. Harvey Lectures. Williams and Wilkins, Baltimore, 1937, p. 169.
(75) Lewis, T. and E. E. Pochin. Effects of asphyxia and pressure on sensory nerves of man. Clinical Science 3: 141, 1938.
(76) Hardy, J. D., H. Goodell and H. G. Wolff. Studies of pain: Observations on the hyperalgesia associated with referred pain. Am. J. Physiol. 133: 316, 1941.
(77) Lehmann, J. E. The effect of asphyxia on mammalian nerve fibres. Am. J. Physiol. 119: 111, 1937.
(78) Wortis, H., M. H. Stein and N. Joliffe. Fiber dissociation in peripheral neuropathy. Arch. Int. Med. 69: 222, 1942.

(79) GORDON, G. AND D. WHITTERIDGE. Conduction time for human pain sensation. Lancet 2: 700 (Dec. 4) 1943.
(80) WOLFF, H. G. Headache and other head pain. Oxford University Press, New York and London, 1947.
(81) WOLF, S. G. AND H. G. WOLFF. Human gastric function. Oxford University Press, New York and London, 1943 and 1947.
(82) CLARK, D., H. B. HOUGH AND H. G. WOLFF. Experimental studies on headache: Observations on headache produced by histamine. Arch. Neurol. and Psychiat. 35: 1054, 1936.
(83) PICKERING, G. W. AND W. HESS. Observations on the mechanism of headache produced by histamine. Clinical Science 1: 77, 1933.
(84) NORTHFIELD, D. W. C. Some observations on headache. Brain 61: 133, 1938.
(85) KARL, R. C., G. E. PEABODY AND H. G. WOLFF. The mechanism of pain in trigeminal neuralgia. Science 102: 12, 1945.
(86) WOLF, S. AND J. D. HARDY. Studies on pain. Observations on pain due to local cooling and on factors involved in the "cold pressor" effect. J. Clin. Investigation 20: 521, 1941.
(87) SCHUMACHER, G. A. AND H. G. WOLFF. Experimental studies on headache. a) Contrast of histamine headache with the headache of migraine and that associated with hypertension. b) Contrast of vascular mechanisms in preheadache and in headache phenomena of migraine. Arch. Neurol. and Psychiat. 45: 199, 1941.
(88) GRAHAM, J. R. AND H. G. WOLFF. Mechanism of migraine headache and action of ergotamine tartrate. A. Res. Nerv. and Ment. Dis. Proc. 18: 638, 1937. Arch. Neurol. and Psychiat. 39: 737, 1938.
(89) SUTHERLAND, A. M. AND H. G. WOLFF. Experimental studies on headache. Further analysis of the mechanism of headache in migraine, hypertension and fever. Arch. Neurol. and Psychiat. 44: 929, 1940.
(90) KUNKLE, E. C., B. S. RAY AND H. G. WOLFF. Studies on headache: The mechanism and significance of the headache associated with brain tumor. Bull. N. Y. Academy of Med. 18: 400, 1942.
(91) KUNKLE, E. C., B. S. RAY AND H. G. WOLFF. Studies on headache: An analysis of the headache associated with changes in intracranial pressure. Arch. Neurol. and Psychiat. 49: 323, 1943.
(92) WOLF, S., H. GOODELL AND H. G. WOLFF. Unpublished observations on comparisons of intensity of pain arising from various sources.
(93) MACKENZIE, J. Symptoms and their interpretation. 3rd ed., London, Shaw and Sons, 1918.
(94) PAYNE, W. W. AND E. P. POULTON. Experiments on visceral sensation: Relation of pain to activity in the human oesophagus. J. Physiol. 63: 217, 1927.
(95) PALMER, W. L. The pain of peptic ulcer. A. Res. Nerv. and Ment. Dis. Proc. 23: 302, 1943.
(96) MCLELLAN, A. AND H. GOODELL. Pain from the bladder, ureter and kidney pelvis. A. Res. Nerv. and Ment. Dis. Proc. 23: 252, 1943.
(97) SIMONS, D. J., E. DAY, H. GOODELL AND H. G. WOLFF. Experimental studies on headache: Muscles of the head and neck as sources of pain. A. Res. Nerv. and Ment. Dis. 23: 228, 1943.
(98) SIMONS, D. J. AND H. G. WOLFF. Studies on headache: Mechanisms of chronic post traumatic headache. Psychomatic Med. 8: 227, 1946.
(99) MCAULIFFE, G. W., H. GOODELL AND H. G. WOLFF. Experimental studies on headache: Pain from the nasal and paranasal structures. A. Res. Nerv. and Ment. Dis. Proc. 23: 185, 1943.
(100) TRAVELL, J. AND N. H. BIGELOW. Referred somatic pain does not follow a simple "segmental" pattern. Federation Proc. 5: 106, 1946.

(101) Travell, J., C. Berry and N. Bigelow. Effects of referred somatic pain on structures in the reference zone. Federation Proc. 3: 49, 1944.
(102) Kunkle, E. C., H. Goodell and H. G. Wolff. Unpublished observations quoted by H. G. Wolff. Harvey Lectures 39: 39, 1943, 44.
(103) Kunkle, E. C. and G. C. Armistead. Unpublished observations.
(104) Ray, B. S. and H. G. Wolff. Studies on pain. "Spread of pain"; evidence on site of spread within the neuraxis of effects of painful stimulation. Arch. Neurol. and Psychiat. 53: 257, 1945.
(105) Gellhorn, E. and L. Thompson. The influence of muscle pain on cortically induced movements. Am. J. Physiol. 142: 231, 1944.
(106) Thompson, L. and E. Gellhorn. The influence of muscle pain on spinal reflexes. Proc. Soc. Exper. Biol. and Med. 85: 105, 1945.
(107) Head, H. On disturbances of sensation with especial reference to the pain of visceral disease. Brain 17: 339, 1894.
(108) Sturge, W. A. The phenomena of angina pectoris, and their bearing upon the theory of counter-irritation. Brain 5: 492, 1883.
(109) Weiss, S. and D. Davis. The significance of the afferent impulses from the skin in the mechanism of visceral pain. Skin infiltration as a useful therapeutic measure. Am. J. Med. Sci. 176: 517, 1928.
(110) Morley, J. Abdominal pain. New York, Wm. Wood and Co., 1931.
(111) Woollard, H. H., J. E. H. Roberts and E. A. Carmichael. An inquiry into referred pain. Lancet 1: 337, 1932.
(112) Wolf, S., H. G. Wolff and H. Goodell. Unpublished observations.
(113) Simons, D. J. Personal communication quoted by Schumacher, G. A., B. S. Ray and H. G. Wolff. Experimental studies on headache: Further analysis of histamine headache and its pain pathways Arch. Neurol. and Psychiat. 44: 701, 1940.
(114) Ross, J. On the segmental distribution of sensory disorders. Brain 10: 333, 1888.
(115) Hardy, J. D., H. G. Wolff and H. Goodell. Observations on pain intensity. Unpublished.
(116) Wolff, H. G. and H. Goodell. Observations on hyperalgesia, see Harvey Lectures 39: 39, 1943, 44.
(117) Gerard, R. W., W. H. Marshall and L. J. Saul. Electrical activity of the cat's brain. Arch. Neurol. and Psychiat. 36: 675, 1936.
(118) Adrian, E. D. Sensory areas of the brain. Lancet 2: 33, 1943.
(119) Head, H. and G. Holmes. Sensory disturbances from cerebral lesions. Brain 34: 102, 1911.
(120) Kendall, D. Some observations on central pain. Brain 62: 253, 1939.
(121) Van Wagenen, W. P. Personal communication to Dr. A. Earl Walker, see reference 31.
(122) Watts, J. W. Personal communication, quoted by Wolff in Harvey Lecture. See reference 59.
(123) Walker, A. E. Personal communication, quoted by Wolff in Harvey Lecture. See reference 59.
(124) Wolf, G. A., Jr. The effect of pain on renal function. A. Res. Nerv. and Ment. Dis. Proc. 23: 358, 1943.
(125) Wolf, G. A., Jr. Mechanism of reflex anuria. Annals Int. Med. 23: 99, 1945.
(126) Gold, H., N. T. Kwit and W. Modell. The effect of extra-cardiac pain on the heart. A. Res. Nerv. and Ment. Dis. Proc. 23: 345, 1943.
(127) Chapman, W. H., M. E. Cohen and S. Cobb. Measurements of levels of heat stimulus perceived as painful and producing wince and withdrawal reactions in patients with neurocirculatory asthenia, anxiety neurosis, or effort syndrome and in control subjects. J. Clin. Investigation (in press) 1946.
(128) Livingston, W. K. Pain mechanisms. A physiologic interpretation of causalgia and its related states. MacMillan Co., New York, 1943.

HEAT: MAN'S EXCHANGES AND PHYSIOLOGICAL RESPONSES

WILLARD MACHLE AND T. F. HATCH

405 Lexington Avenue, New York 17

The purpose of this paper is to summarize the applications of physical laws of heat transfer to analysis of thermal relations between man and his environment; to review briefly the important recent work on the physiological responses of man to heat, and to emphasize some of the studies carried out during the war years which have advanced our knowledge in this field. Systematic exploration of the applications of the laws of heat transfer as they apply to man, is not usual in a review which is dedicated to effects rather than causes. Prior to the war there was some tendency to separation of the fields of investigation. The necessities, however, for recommendation on matters of military usefulness required pursuit of both approaches for decision. Recognizing that present knowledge does not permit fully descriptive mathematical expression for thermal relationships and effects, we have nonetheless used what is at hand and attempted to integrate this into the somewhat disrupted pattern which our present knowledge makes.

A complete bibliography will not be presented for reason that much foreign literature is not available and various reports by military agencies remain classified and not accessible. Adequate general reference lists may be found in earlier publications (1, 2, 3, 4) and in the Annual Reviews of Physiology (1941 to 1946).

Military exigencies largely determined the direction taken by the studies of the war years and have emphasized their importance (5, 6, 7, 8, 9). Though man lives and works in all naturally occurring climates, the equipment in military operations (tanks (10), ships (11, 12), airplanes (13)) and the requirements for special clothing, may create local ambient or personal conditions which are intolerable. This, and the need at all times for maximum efficiency in military personnel, have resulted in the greater bulk of the work of the past five years being directed to the following:

A. Extension to extreme ranges of the more precise work of the Pierce Laboratory, Russel Sage Institute, etc.

B. Further attempts to define the upper level of tolerance to heat.

C. Enlargement of studies on acclimatization, its nature, course and duration.

D. Efforts to devise means for evaluation of total thermal stress from known ambient conditions, and correlation of the physiologic responses of man with total thermal stress.

Because of the large number of variables involved and the requirement for immediate answers to problems of military importance, practically all work has been limited to studies on young, healthy men. Little work has been reported on women, children and the aging (14, 15, 16), with only an occasional reference to the unfit or diseased (17). That differences may be great with respect to both tolerance and acclimatization as the result of individual factors, age, or disease,

is a common observation of all who have worked in the field. The important rôle of age and fitness has been commented on (18, 19, 20, 21, 22, 23, 24, 25). The characteristics of military population, their selection, and fitness, and the requirement for limitation of variables in physiological investigation resulted, therefore, in a great deal of work which is of considerable interest physiologically and directly applicable to young, healthy, male populations engaged in specified types of activity. Extension of results to industrial and demographic problems is limited.

MATHEMATICAL CONSIDERATIONS. The physical laws of heat transfer are usefully employed in the analysis of thermal relations between man and his environment. Most of the recent work in this connection stems from the systematic investigations of the Pierce Laboratory.

It is evident, however, that the complex physiological variability of the human body precludes the possibility of formulating rigid equations which fully describe thermal relations under all conditions. Thus, any equation developed by mathematical analysis must be regarded as an over-simplified statement which may fail to predict behavior in individual cases, not because the physical laws are inapplicable but for the reason that all variables are not properly included. Difficulties in this regard are overcome to some extent in experimental work by a statistical approach, working with large numbers of observations on many subjects. In this way, trend lines are established in spite of the scatter of individual observations, from which useful physical constants are derived.

With proper recognition of its limitations as well as capacities, an understanding of the mathematical approach is essential to a thorough analysis of the problem.

1. *Heat Balance Equation.* Mathematical analysis, as developed at the Pierce Laboratory, starts with the basic equation of heat balance:

$$M + D - V = R + C + E$$

where M = metabolic rate
D = rate of change of body heat content
R, C and E = rates of heat exchange with the environment by radiation, convection and evaporation, respectively[1]
V = rate of heat exchange by respiration

The terms to the right of the equality sign may be re-written, in accordance with the established laws of heat transmission, in terms of the *coefficients* of heat exchange and temperature and vapor pressure gradients:

$$M^1 + D = K_rA(t_s - t_w) + K_cA(t_s - t_a) + K_eA(VP_s - VP_a) \qquad (2)$$

[1]Note that D, R and C may have positive or negative values. D (heat debt) is positive when the body is cooling; R and C are positive when the environmental temperatures are less than skin temperature. Note, also, that these are all *rates* of exchange—heat units per interval of time.

where $M^1 = M - V$
A = body surface area
t_s = average skin temperature
t_w = mean radiant temperature of surroundings
t_a = air temperature
VP_s = vapor pressure of water at temperature t_s
VP_a = vapor pressure of moisture in the air (= per cent R.H. $\times$ VP at t_a)
K_r, K_c and K_e = coefficients of heat exchange by radiation, convection and evaporation.

The temperature and vapor pressure gradients can be determined by direct measurement. The coefficients, however, are not predictable from physical relationships and must be determined experimentally.

2. *Coefficients of Heat Exchange.* A. *Radiation.* The Stefan-Boltzman equation for heat exchange by radiation applies to man as well as to physical bodies. It has the form:[3]

$$R = KA_r(T_s^4 - T_w^4)$$

$$\text{or } R = KA_r[T_s^3 - 6T_s^2(t_s - t_w) + 4T_s(t_s - t_w)^2 - (t_s - t_w)^3](t_s - t_w)$$

where K = universal radiation constant = 4.92×10^{-8} Cal./M²/hr.
A_r = effective radiation area
T_s and T_w = mean radiation temperatures (absolute) of the body surface and surroundings, respectively.

This equation may be simplified to:

$$R = K_rA(t_s - t_w)$$

where $K_r = K(A_r/A)[T_s^3 - 6T_s^2(t_s - t_w) + 4T_s(t_s - t_w)^2 - (t_s - t_w)^3]$
A_r/A = ratio of effective radiation area to total body surface.

Over a moderate range of wall temperatures, K_r remains substantially constant, the theoretical value ($A_r = A$) being as follows for the assumed values of t_w and t_s:

t_w	t_s	K_r
20	32	5.2
30	33	5.7
40	35	5.9
50	38	6.3

The ratio $A_{r/A}$ has been determined experimentally at Pierce Laboratory (26, 27) and at the Armored Medical Research Laboratory (28), with the following reported values:

Sitting, nude and lightly clothed	0.71–0.75
Standing, nude	0.93

Thus, in a hot environment, a standing man (A = 1.8M^2) gains by radiation: $6.3 \times 0.9 \times 1.8 = 10$ Cal./hr. per degree C temperature difference between sur-

[3] Surrounding surfaces are taken to be great compared with man's body area.

roundings and body surface. Heat transfer by radiation is, of course, independent of air movement and, for practical purposes, is not influenced by the moisture content of the air. Against the long wave radiation encountered indoors, ordinary clothing behaves essentially like black regardless of color and, aside from the slight increase in effective radiation area, the effect of clothing upon radiation exchange is negligible. Of the three coefficients of exchange for man, K_r is most firmly established.

B. *Convection.* Heat exchange by convection varies with air movement. Even under still air conditions (no positive air movement, deliberately maintained), natural convection is established around the body and the greater the temperature difference $(t_s - t_a)$, the stronger is the air movement, in proportion to the fourth root of the difference. The effect of positive air movement is to superimpose an added velocity effect which, even with low velocities, exceeds the influence of natural convection, so that in an equation expressing the relation between K_c and v, the natural convection component can be ignored. Thus, the simple expression:

$$K_c = K_c^1\sqrt{v}$$

has been found adequate to describe the relationship over a wide range of velocities (15 fpm to 600 fpm). It will be noted that when $v = 0$ in this equation, $K_c = 0$; the velocity, however, is the absolute value and includes natural convection so that, in practice there is no such thing as zero velocity, and heat exchange by convection is never absent (except when $t_s - t_a = 0$). Experimental determinations of K_c have shown that it is highly variable; it varies with body build, the position of the body and for the same measured velocity it will take different values depending upon the direction and complexity of air movement. Under still air conditions values differ because of movement of the subject and variability of natural convection currents. Working with complex vertical air movement around the sitting subject, the following relationship was established at the Pierce Laboratory (27):

$$K_c = 1.04\sqrt{v} \text{ (cm./sec.)}$$

The velocity, in this case, was measured in the empty booth since inconsistent values were obtained when readings were taken in the occupied chamber. At Fort Knox, in a wind tunnel with linear horizontal flow, and with subjects standing nude, the relationship was found to be (28):

$$K_c = 0.74\sqrt{v} \text{ (cm./sec.)}$$

Lack of agreement between the two is accounted for by differences in test arrangements.

The rate of heat loss by convection per unit of body area from light clothing was found not to differ greatly from that of the nude man, despite the greater area of the clothing. This does not hold true, of course, in the case of bulky winter clothing.

In round numbers, a convection coefficient of 2.0 for still air was found by the

Pierce Laboratory. This is about one-half the value of the radiation coefficient. Thus, under comfortable conditions with 25 per cent evaporative heat loss, losses by radiation and convection would be 50 per cent and 25 per cent, respectively. At higher air velocities, however, the proportions are different. At an air velocity of 100 fpm, K_c is approximately 5, and the partition of losses would be: $E = 25$ per cent; $R = 32$ per cent; $C = 43$ per cent.

Compared with K_r, the value of K_c is less well established, although the mathematical form of its relation to air velocity can be accepted.

C. *Evaporation.* Physically, the rate of evaporation from a wet surface varies with the area and shape of the surface, with the vapor pressure gradient and with air velocity. Various empirical and semitheoretical equations have been developed which are helpful in considering human evaporation, but only to a limited extent. These physical equations are for surfaces which are completely and continuously wet—a situation which apparently never occurs in the case of man. As a consequence, the term K_e contains a variable factor of wetted area as well as the others properly included in the coefficient. The degree of wetness varies from a minimum under conditions of normal comfort when no active sweating takes place, up to a maximum value in extreme heat when the sweat rate exceeds the evaporative capacity.

1. Maximum evaporative capacity. The mathematical significance of the evaporative mechanism was developed by Gagge (29), who determined experimentally an apparent maximum value of approximately $K_e = 3.0$ at a single air velocity (8.5 cm/sec., vertical) for sitting subjects (nude) in a moderately severe environment. More extensive data obtained at A.M.R.L., Fort Knox (28) on standing nude men in a linear horizontal air stream gave the following relation between K_e and v for conditions of maximum evaporation:

$$K_e = 1.84\, v^{0.37} \ (v \text{ in cm./sec.})$$

At 8.5 cm./sec., the Pierce Laboratory velocity, this gives $K_e = 4.05$, or some 30 per cent greater than the value reported by Gagge. Differences in test arrangements as well as more severe environments undoubtedly account for the higher Fort Knox figure.

The exponential form of the equation is in agreement with the relationships developed for physical bodies such as spheres and cylinders. It is of interest that the coefficient for man is considerably below the values for simple physical bodies, suggesting that even under conditions of maximum sweating, only a fraction of the body area is utilized in evaporation.

2. Less than maximum evaporation. In environments not sufficiently severe to induce maximum evaporation, the sweat rate is adjusted physiologically to an amount which satisfies the heat balance equation. Evaporative heat loss is then given by the equation:

$$E = K_e^1 A(VP_s - VP_a)$$

where $K_e^1 = K_e \dfrac{A_e}{A}$

$\dfrac{A_e}{A}$ = effective evaporative surface, relative to maximum.

The process of adjustment is complex and does not lend itself to direct mathematical description. An increase in skin temperature accompanies the increased sweating rate and thus reduces the heat gain by radiation and convection, which, in turn, lessens the requirement for evaporation. This mutual adjustment continues until a kind of physiological balance is obtained. Under equilibrium conditions, we may write:

$$\frac{A_e}{A} = \frac{M' - R - C}{E \text{ max.}} = \frac{M' - K_r A(t_s - t_w) - K_c A(t_s - t_a)}{K_e A(VP_s - VP_a)}$$

Empirically, $VP_s = 2.5\, t_s - 45$,[3] from which:

$$\frac{A_e}{A} = \frac{M' - K_r A(t_s - t_w) - K_c A(t_s - t_a)}{K_e A(2.5\, t_s - VP_a - 45)} \tag{3}$$

It is evident from this equation that $\frac{A_e}{A}$ is a function of skin temperature as well as of the environmental conditions. As a consequence, the value of $\frac{A_e}{A}$ cannot be calculated from this relationship for a given environment without an independent means of predicting t_s. It is true that t_s does not vary greatly, but even a difference of 1°C in an assumed value of t_s would cause a considerable change in the ratio. For example, assume an environment of $t_a = t_w = 40$°C; $VP_a = 13$ mm; $v = 10$ cm./sec., ($K_r = 5.7$; $K_c = 2.3$; $K_e = 4.3$). For $t_s = 34$°C, $A_e/A = 0.85$. A 1° increase to $t_s = 35$°, however, decreases A_e/A to 0.71, according to this equation. No mathematical means has yet been devised for solving this dilemma.

3. Upper limits of evaporative capacity. In view of the maximum value of K_e, it has been suggested that there are limiting combinations of temperature and humidity (for any given air velocity) for evaporative control beyond which excessive body heating occurs. Using an evaporative coefficient of 3.0 and an assumed constant skin temperature of 35.6°C, such limiting values of dry and wet bulb temperature were calculated (30). It has been found, however, that the skin temperature is not constant under all combinations of thermal stress which elicit maximum evaporation. Hence, the lines thus calculated will not state the case correctly. At $t_a = t_w = 49$°C and $v = 15$ cm./sec. ($K_c = 8.0$; $K_e = 5.4$), for example, we find for $t_s = 35.6$°C that the limiting $VP_a = 11$ mm.; taking $t_s = 37$°C, on the other hand, increases the limiting value of VP_a to 16.9 mm. Thus, by assuming only slightly higher skin temperature, the acceptable moisture content of the air has increased 50 per cent. It is evident that until an independent means of predicting t_s is developed, such calculations are of limited value.

3. *Equilibrium State.* Provision is made in the basic heat balance equation for any changes which occur in body heat content, so that the equation is applicable at all times, whether the subject is in a steady or changing thermal state.

[3] The simple linear equation holds with little error over limited ranges of temperature, such as the range in variability of skin temperature considered here.

When man passes from a normal to a hotter environment, he heats up. Because of his considerable heat capacity and highly flexible sweating mechanism, the resulting increase in body temperature is commonly not great and may escape attention. Nevertheless, there is an increase in body temperature which continues until a new equilibrium is established (provided the subject remains in the environment and his normal physiological processes of thermal adjustment are not altered). Obviously, it is important to know how long a period is required to approximate the new equilibrium state.

The change in body heat content can be written:

$$D = SW(a\Delta t_r + b\Delta t_s)$$

where S = specific heat of body
W = body weight
Δt_r and Δt_s = *rates* of change of internal and skin temperature.
a and b = weighting factors which proportion body heat in relation to internal and skin temperatures.

In terms of the calculus, the heat balance equation can be written:

$$-\frac{SW}{A}(a\,dt_r + b\,dts) = \left[M' + K_r t_w + K_c t_a + K_e\frac{A_e}{A}(45 + VP_a) - (K_r + K_c + 2.5\,K_e)t_s\right] d\theta$$

where dt_r, dt_s = differential changes of body and skin temperature.
$d\theta$ = differential time interval.

Body and skin temperatures are mutually dependent, and, over a limited range we may write: $t_r = m + n\,t_s$, from which $dt_r = n\,dt_s$.

In a given situation, the terms: $[M^1 + K_r t_w + K_c t_a + K_e\,(45 + VP_a)]$ and $(K_r + K_c + 2.5K_e)$ are constant. Calling these C and C^1, respectively, we may write:

$$-\frac{SW}{A}(an + b)\int_{t_0}^{t_s}\frac{dt_s}{C - C'dt_s} = \int_0^{\infty} d\theta$$

which gives

$$\frac{t_e - t_s}{t_e - t_o} = e^{-\frac{AC'}{SW(an+b)}\theta}$$

where t_o = initial skin temperature
t_e = new equilibrium skin temperature
t_s = skin temperature at time, θ.

This equation has the form of the common law of equilibrium and with it, the time required to reach any given percentage of equilibrium can be calculated, provided the values of the several factors are known. It will be noted that the relationship is independent of the environmental temperature, that is, the time required to make 50 per cent (e.g.) of the adjustment to the new environment is the same regardless of the magnitude of shift from one environment to the other.

The rate of body heat production also has no direct influence[4]. On the other hand, the significant factors are: 1, pertaining to the environment: the coefficients of heat transfer, K_r, K_c and K_e, and 2, pertaining to man: heat capacity per unit area, $\frac{SW}{A}$ and circulatory indices, a, b and n. The relationship is comparable to the process of adjustment to an atmospheric contaminant, such as benzol. Continued exposure to the benzol vapor results in a building up of concentration of the contaminant in the body, the rate of increase becoming less as equilibrium is approached. The ultimate body concentration is directly proportional to the atmospheric concentration but the time required to make the adjustment is the same with 1000 ppm in the air as with 10 ppm. The rate is dependent upon the volume rate of breathing and blood solubility (comparable to K_r, K_c and K_e), upon blood circulation rate (compared with a, b and n) and body weight (heat capacity).

Data are limited with respect to the values of a, b and n. For a limited range of exposures for resting subjects, a and b have the reported values: $\frac{2}{3}$ and $\frac{1}{3}$ respectively (31, 32). From A.M.R.L. data, n for standing subjects exposed to heat may be taken as $\frac{1}{3}$. Using these values and assuming an average man: $W = 70\ Kg$, $A = 1.8\ M^2$, $S = 0.82$, we find, for an air velocity of 30 fpm ($K_r = 5.7$, $K_c = 2.9$, $K_e = 5.0$, and $C^1 = 21.1$), the value of the exponent to be $-1.18\ \theta$. Thus, the time required for 50 per cent adjustment is 0.59 hour = 35 minutes and 2 hours—33 minutes' exposure is necessary to attain 95 per cent of the temperature rise to the new equilibrium. At a higher velocity, say 600 fpm ($K_r = 5.7$; $K_c = 13.0$ and $K_e = 15.4$; $D = 58.0$), the corresponding times are: 13 minutes and 55 minutes.

Equilibrium adjustment, as delineated by the above equation is, of course, oversimplified, and ignores certain physiological aspects of the problem to a considerable degree, as will be seen from later discussion. Considering its physical basis, however, it is evidently a close approximation of the true situation, and despite the limited knowledge of the magnitude of certain factors in the equation, especially a, b and n, and the variability of others, the predicted values in the foregoing example are in approximate agreement with experience. Thus, in one of the Fort Knox experiments (33), the temperature rise curves are found to be exponential, in accordance with the equation and to approach equilibrium at the same rate for all environmental conditions, varying from moderate to extreme. These curves yield an average exponent of $1.08\ \theta$, which compares well with the values calculated above.

The time rate of approach to equilibrium is an important factor in designing experiments for the quantitative study of thermal physiology and in the interpretation of results. Experiments of short duration are limited to the steep part of the equilibrium curve where the change in body heat content, D, which can be determined only indirectly, has the greatest value. Observations are the least stable, not only for the body as a whole, but more particularly, from one part to

[4] The ultimate equilibrium level t_s, toward which the subject is tending is a function of the metabolic rate.

another. To illustrate: the specific heat per unit area is much smaller for a leg or arm than for the torso. Thus, other things being equal, the rate of adjustment in the torso is slower. Readings are subject to relatively greater error than under steady-state conditions and may even lead to erroneous conclusions respecting the relative influence of different factors. For example, consider two physically equivalent environments with respect to the combined effects of radiation, convection and evaporation, such that they produce equal skin temperatures at equilibrium, one with relatively still air ($K_r = 5.7, K_c = 2.5, K_e = 4.4$) and the other with high velocity ($K_r = 5.7, K_c = 13.0, K_e = 15.6$). The respective values of the exponent in the equilibrium equation would be (for the average man): $-1.08\,\theta$ and $-3.3\,\theta$. Hence, at the end of one hour of exposure, under the still air conditions the skin temperature will have increased only 65 per cent of the ultimate rise whereas with the high wind velocity in the second situation, 96 per cent of the adjustment will have occurred. For a 5° maximum increase, there would be a 2° difference in t_s at the end of one hour exposure, from which it might be erroneously concluded that the two environments were not equivalent.

The concept of equilibrium is important in the consideration of upper limits of tolerance to heat. The statement is frequently made that a given environment. is intolerable because it imposes a thermal demand beyond man's capacity of adjustment and, hence, his ability to maintain a constant body temperature By implication, and often directly, it is concluded that the body temperature will continue to rise so long as the subject remains in the environment. Such a conclusion is erroneous (assuming that maximum sweating rate, circulatory capacity and other physiological limitations are not exceeded). The process of heating up is a temporary one and continues only until a higher skin temperature is reached where E balances $M + R + C$. The equilibrium phenomenon has been demonstrated even in environments well beyond the practical physiological limit of tolerance, involving marked increases in both internal and skin temperatures (33).

4. *Equivalent Environments.* An aspect of thermal physiology of very considerable practical as well as fundamental interest has to do with the question of equivalence of environments. Since the three environmental avenues of heat exchange, radiation, convection and evaporation, are capable of independent variation, it is evident that different combinations can have the same summation effect and thus be equivalent. At the outset, however, in an analysis of equivalence, it is necessary to define the criterion of equality. This has been done in various ways, as: subjective, in terms of relative comfort determined by test subjects (e.g., the "Effective Temperature" of the A.S.H.V.E.); physiological, by an empirical combination of physiological indices (23); physical, derived in various ways from the mathematical equations of heat exchange (30, 34). Each has its own particular advantages and limitations and agreement between them in respect to equivalence of environments is not necessarily expected, owing to basic differences in the initial criteria of equality. Difficulty arises from the fact that two different sets of factors must be combined; not only must the proper

combination of the physical environmental factors be discovered, but also, it is necessary to combine properly a group of complex physiological factors which, together, determine the total strain and which, at the present time cannot be described in absolute terms.

The present discussion will be limited to indices having mathematical origin.

A. *Operative temperature.* Heat exchange by radiation and convection, given by the equation: $R + C = K_r A(t_s - t_w) + K_c A(T_s - t_a)$, may be re-written (34):

$$\frac{R + C}{A} = K_r + K_c\left[t_s - \frac{K_r t_w + k_c t_a}{K_r + K_c}\right] = K_o(t_s - t_o)$$

where

$$K_o = K_r + K_c$$

$$t_o = \frac{K_r t_r + K_c t_a}{K_r + K_c}$$

The combined temperature, t_o, called the Operative Temperature, properly weights the effects of the two elements, R and C, according to their respective coefficients and gradients. Thus, all environments having the same t_o are said to be equivalent in that they impose the same caloric demand by radiation and convection. It is to be noted, however, that this is true only if K_c remains the same. For other K_c values, given by different air velocities, the effect upon skin temperature will be different, necessitating further compensation in environmental temperatures to keep t_s and $R + C$ the same.

B. *Equivalence in terms of skin temperature.* Returning to equation (3), and solving for t_s,

$$t_s = \frac{M'/A + K_r t_w + K_c t_a + K_e \frac{A_e}{A}(VP_a + 45)}{\left(2.5K_e \frac{A_e}{A} + K_r + K_c\right)} \tag{4}$$

By definition, one could say that all environments giving equal t_s values by this equation are equivalent. It will be noted, however, that the value of $A_{e/A}$ is not definitely known, so that the equation has little practical value until another relationship between $A_{e/A}$ or t_s and environmental stress is developed. In general, it appears that skin temperature increases with air and wall temperatures but for any given combination of temperature and atmospheric moisture content, t_s decreases with increasing air velocity. Thus, it is evident that there is no direct correlation with evaporative demand and sweat output. $\frac{A_e}{A}$ also decreases with increasing air velocity in any given environment below those which elicit maximum evaporation, in accordance with the equation

$$\frac{A_e}{A} = cv^d$$

where c and d apparently vary with environmental temperature and moisture content (or the gradients of temperature and vapor pressure). These indications emerge from Nelson's data (28), which, however, are not extensive enough to permit the determination of the actual values of c and d in relation to the thermal stress.

5. *Physical Instruments for Measuring Equivalence.* Many attempts have been made to develop a single instrument which responds thermally to the environment in a manner corresponding to the human body, thus providing a single equivalent thermal index of the environment in place of the four separate factors: radiant temperature, air temperature, moisture content and air movement. All such attempts have failed for several basic reasons: first, the coefficients of heat exchange of the instrument have not been like those of man, but more importantly, no provisions have been made for equivalent skin temperature variation, which involves in the human body changes in internal conductance and in sweating rate and distribution. With the body coefficients of heat exchange known, it would not be difficult to design an instrument having the same coefficients, but it is not likely that the complex variability in internal conductance of the body and evaporation rate can be duplicated. Thus, it appears doubtful if the objectives of such an instrument can be met. A physical instrument with the proper coefficients and maintained at a selected *average* surface temperature would, however, provide a useful index of the cooling power of the environment although it does not appear that it would provide any basic information beyond that given by separate measurements of t_w, t_a, v and VP_a.

PHYSIOLOGICAL EFFECTS OF HEAT. The heat balance equation given above describes a state. Physiological disturbances in the organism result from increase in body heat content with concomitant temperature rise, and from operation of the homeostatic mechanisms of temperature regulation. Most studies have dealt with measurement of the basic change, i.e., increased internal (rectal) temperature and with the associated effects upon the circulation and volume of sweating, since these are convenient and easily measurable indices of the degree of strain upon the organism. Further evaluation of the physiological effects may be made by means of the appearance and behavior of men, the subjective state, functional efficiency, and the occurrence of heat exhaustion (18, 19, 21, 25, 33, 35, 36, 37, 38, 39, 40). In the interpretation of any observations it is important to keep in mind that the basic physiologic change is an increase of the body heat content associated with the new equilibrium level, and that the capacity of the individual to perform efficiently is the only definitive criterion of overall effect, most other phenomena being evidences of load upon the homeostatic mechanisms. The lack of ability to predict performance on the basis of evidence of circulatory system strain or increase in body heat content alone, has been emphasized (21, 25, 41, 42, 43).

1. *Body Temperature and Gradients.* Since the time of Claude Bernard (44), the temperature of the internal mass of the body has been the subject of speculation and study, and it is recognized that there are local variations in general order and direction of the fluctuations in temperature which change with ac-

tivity, alterations in local circulation, ambient temperatures, time, level of thermal equilibrium, etc. (1, 3, 45, 46). The rectal temperature properly measured has been accepted by practically all workers as the arbitrary index of internal body temperature. Spealman (47), however, believes that temperatures taken within the stomach are possibly a more reliable index of deep body temperature than measurements made within the rectum, because the close proximity of surrounding cold water to the junction of the rectal thermocouple may cause falsely low readings. Our own observations do not support this.

It is to be expected that lags will occur in any changing state. The temperature of the internal organs responds somewhat more promptly to heating and cooling than does the rectal temperature, the temperature of the viscera reflecting the changes in heat production of these organs, and significant thermogenesis occurs in the intestine of dogs (48, 49). In thermoelectric measurements of blood in the aorta, portal and hepatic veins, the lowest temperature was found in the aorta and the highest in the hepatic vein. That the viscera are largely responsible for heat production and that resting muscle contributes little is pointed out by Milhorat et al. (50), who compared results from 50 normal individuals and 100 patients with muscular diseases. In pre-pubertal patients muscular wasting up to 60 per cent had no effect upon heat production and muscle wasting in adults was disproportionally great when compared to slight reductions in heat production that were observed. Barcroft and Edholm (51) found temperatures of from 18 to 44 degrees C in the brachioradialis muscle after immersion of the arm in water from 12 to 45°C. They point out the various factors which determine local temperatures. All of the recent work and earlier analysis (1) lead one to the inference that the concept of a gradient may imply a constancy of conditions which can exist only in highly controlled and constant situations. The thermogeography of the body had perhaps best be looked upon as a series of constantly shifting regional equilibria or exchanges, with differing levels subject to the variables operating at the moment. Considerably more data will be needed before the pattern and order of magnitude of the subsurface thermal shifts can be synthesized.

The frequent employment of skin temperature data in studies of thermal exchange has provided a large volume of new observations, much of which has been integrated into other studies (23, 28, 33, 41, 52, 53, 54). The most satisfactory tool for measurement remains the radiometer as described by Hardy and DuBois (55). As pointed out above, the use of skin temperature alone has limited value in interpretation of evaporative loss.

2. *Peripheral Blood Flow; Vasomotor Regulation.* Earlier definitive studies (56, 57, 58) have established the rôle of dilatation of blood vessels in increasing heat elimination and describe the general distribution of vasomotor changes. Additional comprehensive studies by Sheard (59) and other workers (60, 61, 62, 63) have further clarified the regional aspects of the vasodilatation at temperatures ranging from 18° C to 34°C. They found (as have others) that the regulation of heat dissipation at comfortable environmental temperatures (23 to 30°C) is accomplished chiefly by changes in the tonus of cutaneous vessels in the ex-

tremities. At 18°C the superficial vessels of the head and trunk are already near maximal dilatation while those of the arms and legs still show considerable constriction—the legs more than the arms. Between 18 and 25°C, increase in heat dissipation is brought about by dilatation of the vessels of the arms and hands, and at high temperatures (25 to 30°C) additional loss is accomplished by dilatation of the vessels of the skin of the feet and lower legs. Maximum vasodilatation was maintained at ambient temperatures above 31 to 32°C, evaporative heat loss then meeting additional need for dissipation. In contrast to cooling rates of skin, which were exponential, two rates for warming were obtained, consequent upon the effects of both increasing input (vasodilatation), and environmental gain. Evidence is offered that the arterioles of the skin are capable of dilating as a result of direct effects of heat upon the vessel (63). When conditions are extreme, the central mechanism functions as an integrating agency (as it does in the general response to local heating). This central reaction is mediated by capillary dilator fibres of the sympathetic system.

Other studies (64) have pointed out the ease of producing indirect vasodilatation by heat.

The time required for circulation in the forearm to reach equilibrium at various temperatures is indicated by the studies of Barcroft and Edholm (51) who found at 45°C, that the flow increased to a maximum in from 30 to 45 minutes and remained constant thereafter. These findings are very pertinent in the light of the many observations on the time required to reach a steady state of thermal equilibrium. It would be very interesting to know whether this time for maximum dilatation is shortened by repeated exposure or general acclimatization. The magnitude of the increase in peripheral flow has been investigated (65). Using a combination plethysmograph and calorimeter embracing the hand, flow rates of the order of 30 ml./100 ml. tissue in the hand were found at 38°C as compared with volume flows of as little as 0.15 ml./100 ml. tissue with exposure to cold. The maximum rates of flow were lower than some reported (66, 67, 68).

The results of many workers have established that local temperature changes will incite like general and contralateral responses and heat acts as a coronary vasodilator, the effects being reflex in origin (69). That cutaneous vasodilatation is associated with corresponding changes in muscle is demonstrated (51).

3. *Circulation.* Beginning with exposure and gain in heat storage evidence of strain upon the homeostatic mechanisms becomes evident and there is superficial vasodilatation with increased cutaneous circulation. To compensate for the higher rates of peripheral flow, cardiac rate and output increase, blood pressure rises, and sweating begins (70). With extreme ambient conditions and when the heat dissipating mechanism is unable to maintain the body temperature within the limit that is critical for the individual, physiological breakdown occurs. The manifestations are diverse, the outstanding exhibit being inadequacy of circulation with a forward type of failure. Taylor, Henschel and Keys (71) reported results of 7,000 observations on pulse, blood pressure, rectal temperature and rate of sweating of 43 subjects—before and during exposure to dry heat for from 2–8 days. Modified Crampton scores were calculated and the devia-

tions from control values noted. Bean and Eichna (21, 25, 72) have described the cardio-vascular instability in detail and elsewhere Eichna et al. (33) present in tabular and graphic form observations on rectal temperature, heart rate, sweating and skin temperature for a group of 13 experimental subjects exposed to environments ranging from 92.4 to 96. T_w(T_d 92 to 120°F.). The point is made that changes in the circulation do not always parallel or indicate the capacity for overall performance. They noted in their study of more difficult environments (T_w 94 and T_d 94 to T_w 91 and T_d 119.3) that some of their subjects experienced an adjustment like "second wind". During the first hour men worked with difficulty and had heart rates that tended to be higher than expected (an increase of about 25 beats/degree C rise in rectal temperature), while toward the end of the second hour men improved and were in better condition. Decrease in heart rate, despite a rising rectal temperature, occurred in some men in the second hour. Disturbances were most pronounced during the changing state—subjects achieving thermal equilibria for the test environments in about two hours.

A. *Instability of circulation.* Associated with the very considerable shifts in circulation a degree of instability becomes evident. When the stress is great, or after some hours of exposure, especially with work, instability may become pronounced and its effects disabling (21, 23, 25, 71, 72). The instability becomes especially apparent in the upright position, a result of inability to achieve postural adjustment. The circulatory effects of the vertical stance have received much study and have been reviewed (73), and recent work under ordinary thermal conditions has been reported (74, 75). Allen, Taylor and Hall (76) found that 20 per cent of 111 young men developed orthostatic insufficiency on the tiltboard at normal temperatures. These rates are like those of Eichna's (72) group of 150 men, 23 per cent of whom had orthostatic hypotension with syncope on initial exposure to heat. An additional "abnormal" group (21 per cent) included subjects who could remain erect for 3 minutes but whose blood pressure was below 100 mm. Hg and at least 10 mm. below levels under comparable conditions in the cool. After 5 days of work in the heat, only 1.0 per cent of men fainted when tilted and 13 per cent fell into the normal group. Thus, 5 days of repeated exposure and training reduced the percentage of men with some degree of circulatory system inadequacy from 44 per cent to 14 per cent. Allen (76) likewise found that training developed the ability to withstand gravitational stress. Interestingly enough he found that this capacity was not related to the ability to meet the stress of vigorous exercise. It would be helpful to have more information upon the relationships between the ability to withstand orthostatic stress in the cool (with and without work) and the capacity to maintain adequate circulation in the heat.

The reasons for the poor return are not clear at this time. It is known that there is increase in the capillary bed with venous enlargement and relaxation and arteriolar dilatation is perhaps a critical factor (77). That venous pooling with reduction in venous return does occur seems inescapable, and additional data on the magnitude of reduction in venous pressure produced by trapping of the blood in the extremities have been reported (78). The ability to prevent syncope in the

erect position by initiation of limited muscular activity in the legs, the application of cuffs, and the common observation that elevation of the legs will prevent the fall in blood pressure, are all in support of the thesis that circulatory inadequacy is associated with pooling of blood in the extremities and deficiency in venous return. At high temperatures the disturbances, however, comprise much more than a simple disparity between the capacity of the vascular bed and the available volume of blood. We have, unfortunately, no electromyographic studies in heat. Warmth relaxes voluntary muscle and one suspects that lessening in muscle tone with reduction in pumping action (79) may contribute significantly to failure of adequate venous return.

Added loads other than posture have been investigated in relation to thermal stress and numerous reports have pointed out the deteriorating effects of heat in the presence of shock-like states (80, 81, 82). Whether induced by hemorrhage or burns, high environmental temperature has a decisive influence in increasing mortality. In this connection, the disturbances in thermal regulation that occur in congestive heart failure are of great interest. Heat production may be increased in heart failure, but observations in cool climate (83, 84, 85) suggest that the high body temperature results from disturbances in circulation to the skin with resulting impairment of heat loss.

B. *Adequacy of circulation.* The adequacy of circulation is contingent upon a number of factors, important among which are the relationship between the blood volume and the capacity of the vascular system, and normal functional ability to increase cardiac output by increase in rate and stroke volume. Even with adequate venous return, relationships between pulse rate and cardiac output are complicated and with exposure to high temperatures with great demand for peripheral flow and additional inadequacy in venous return, it becomes impossible to evaluate the significance of individual physiological changes as isolated phenomena. The initial and often considerable elevation in pulse rates can give quite an erroneous notion as to cardiac output and adequacy. Pulse rates per degree increase in rectal temperature vary rather widely (from 20–30/min.) but are considerable. Pulse rate and blood pressure are indices of the strain on the circulation rather than of the ability to control rise in body temperature. The general relationship is obvious, but not sufficiently close to be of value in individual instances. As Krogh (86) has observed, with an inadequately filled heart an increase in pulse rate from 75 to 120 would result in a decrease in cardiac output from 3.1 to 2.9 liters. These factors and the marked influence of muscular activity upon venous return, account for the differing statements as to the effects of heat on cardiac output. Wilkins, Hunt and Friedland (87) reported only minor changes in cardiac output with warming and cooling, which was attributed to the buffering capacities of the circulatory system. Earlier observations have established that with initial exposure, at least, there is some increase in cardiac output. Asmussen (88), working with acclimatized subjects, found no significant differences in output in the first half hour of work, after which there was decrease in output, more marked in the hot climate than in the cool. As work was continued and breakdown began, a reduction in stroke volume of 20 per cent associated with very high pulse rates was noted.

Glickman et al. (89) reported results on studies of blood volume as affected by heat. Increase in plasma was noted in 6 experiments, decreases in 4 and no change in 14. Increases that were noted ranged from 5 to 13 per cent. In instances where there was no change in plasma volume the red blood cell volume increased. Reports on blood volume changes continue to be confusing. With respect to Glickman's observations, it may be noted that exposures were quite short (a few hours) and hydration was not maintained. When adequate time is allowed blood volume increases up to 20 per cent or more may be observed (53).

One can think of few things which do not affect the blood pressure—high ambient temperature is no exception. In general, significant deviations are not observed except during the changing state and during the period of instability in the first few days of exposure to high temperatures when there is a tendency to low values especially in the erect stance and precipitate drops may occur before the syncopal attacks. The fall may be in the systolic, diastolic, or pulse pressures (21, 24, 25, 72).

4. *Sweating*. The subject of insensible water loss has been reviewed. Little new has been added since by way of basic concept. Barbour (91), in summarizing his studies of water movement in response to environmental temperature, emphasizes the heat-conserving properties of vapor pressure changes as they affect the availability of fluid for insensible loss and sweating.

Pinson (92), in study of water loss in skin areas in which the sweat glands had been inactivated, found that general exposure to heat caused increases in water loss, insensible perspiration being doubled by an increase in skin temperature of 10°C. Rate of blood flow acted only as it increased the skin temperature.

The variations in rate of sweating in different parts of the body have been studied under subtropical conditions (93). The most rapid rates of insensible loss were from the hands and feet, forehead, and cheeks, the loss being relatively slow from the trunk, arms, and legs. Marked variations from area to area occurred. In experiments carried out under extreme conditions very high rates of sweat loss have been reported (23, 33, 41, 52). Under severe conditions a sweat loss of 2.25 L/hour was common (33). The magnitude of the fluid exchange is apparent from these estimates. With an average man having 3.5 L of plasma (of which 3.2 L are free water), men lost an amount approximating 70 per cent of the free water of the blood. In some instances the subjects lost from 3.3 to 3.9 L of sweat an hour, representing amounts equalling or exceeding (120 per cent) the total blood water. Under the experimental conditions the maximum sweating rate was 4.2 L/hour and some men completed 4 hours of work while sweating at rates of 3.0 L/hour (amounting to 4 times the blood water and about $\frac{3}{4}$ the total extracellular fluids). The significance of these high fluid losses will not be considered here in relation to the problem of fluid balance, save to say that with adequate replacement there were only small changes in cellular volume and in the concentration of electrolytes, decreases in serum sodium, potassium and chloride concentrations being less than 5 m. Eq./L (33). The rôle of evaporative loss has been commented upon above and the problems of relating coefficient of evaporation, skin and air temperature and wind velocity pointed out. The mechanisms for control of sweating have been studied (94),

and the ability of denervated sweat glands to respond directly to heat and mecholyl demonstrated (95). Remote responses are however dependent upon the intact state of the motor side of the reflex arc. The sensory side was found not to be necessary since general sweating followed the action of heated blood upon the central nervous system.

Though the limiting capacity of the circulatory system is usually the determining factor, failure of the sweating mechanism may limit adjustment to heat. Collings (96) found cessation of sweating a prodroma in 26 per cent of cases of heat exhaustion in steel mill workers. Others (40, 97, 98, 99, 158) describe cases of failure of sweating mechanism designated as "thermogenic anhydrosis". The phenomenon was seen after prolonged exposure to desert climates; or it followed exposure to hot humid heat for 7 months (98). The regional distribution of the anhydrosis was similar among the cases and like that which follows thoracolumbar sympathectomy. Since sudomotor, pilomotor and vasomotor responses may be governed by the end organ, sweat gland, vessel, or pilomotor muscle, by reflex through the spinal cord, or by central hypothalamic control, the location of the dysfunction was sought (99). Response to heat, pilocarpine and mecholyl was one of complete anhydrosis in the affected area, with hyperhydrosis above. Recovery was prompt in the cool, when repeat test elicited normal response. Despite the fact that cholinergic drugs failed to evoke sweating, it was not believed that the end plates were the site of involvement, since mecholyl, etc., fails to stimulate sweating in cerebral lesions, heat stroke and thyroid crises. Action is viewed as an active cerebral inhibition which is sufficiently strong to abrogate drug effects. A refractory state of the sweat gland itself has not been excluded (151, 152, 153).

The disastrous effects of heat on patients having hereditary ectodermal dysplasia of the anhydrotic type is well known. Studies of the effects of heat on 2 of 3 brothers without sweat glands have been reported (100). Regional variations in water loss were further studied in diseased subjects (101, 157). The existence of great variation in water loss in the absence of disease was evidenced by one subject, clinically well, but who complained of undue fatigue when working in hot weather. Skin water losses under comfortable conditions (75°F., R.H. 50 per cent) were consistently below average normal, the difference being markedly accentuated in hot, humid conditions (105 T_d, 75 per cent R.H.). A more recent contribution (17) demonstrates that the deficiency in circulation associated with congestive heart failure, results in a reduction in sweating to from 27 to 47 per cent below normal levels for the patient during the compensated state. The deficiency was general, and even mild thermal stress was poorly tolerated by these patients.

5. *Salt Depletion.* Questions of salt requirement, methods of maintaining electrolyte and water balance and the effects of water deprivation will not be considered in this report except for mention where they have a direct bearing upon the effects produced by heat.

That high rates of sweating result in the loss of large amounts of salt and lead to a deficiency state with characteristic disturbances has been well established

(102), and recently, the salt deficiency states have been excellently portrayed (11). The important relationship of salt intake to cardiovascular function and heat exhaustion received limited attention until the investigations of Taylor, Henschel and Keys (41). In a study of 34 subjects, they demonstrated that a moderate salt intake is essential for the maintenance of cardiovascular function and the prevention of heat exhaustion. Men on low salt intake (6 grams) had higher pulse rates and rectal temperatures than did men on normal (13–17 grams) intakes. A significant observation was that salt deficiency with reduction in plasma chloride was not always associated with heat cramps but in many instances led to heat exhaustion. Failure to replace water leads to similar phenomena of dehydration exhaustion (18, 19).

Recent opinion as to salt loss during exposure to heat is summarized in a report of the Council of Pharmacy and Chemistry (103).

Studies of Pitts et al. (104) led to results like those of Taylor, and indicate that while water replacement was best on an hour to hour basis, special circumstances are required to justify salting of the water. They suggest that the chloride concentration of sweat is affected by 3 primary factors: 1, a peripheral factor correlated with skin temperature; 2, a central one correlated with rectal temperature and rate of sweating, and 3, personal idiosyncrasy. They found the rate of sweating and salt content to increase with increasing rectal temperature. Taylor et al. (41) found no correlation, and Horvath and Shelley (52) observed increasing rates of sweating with unchanged rectal temperatures. Consistent differences between subjects have been observed by all investigators. Morreira, Johnson, Forbes and Consolazio (105) found that chloride changes followed the foregoing pattern of responses (104) and that sodium was present in concentrations almost equivalent to chloride and showed the same correlations with skin and rectal temperatures and with rates of sweating. In contrast with sodium, sweat potassium invariably decreased as work was prolonged and showed no correlation with the factors listed above.

6. *Metabolism and Nutrition.* Studies of resting metabolism in the heat are complicated by many factors and results are far from consistent; elevations, depressions, and no change having been reported. Too many studies neglected the factor of acclimatization and the level of thermal equilibrium at the time. There appears to be little question that during the changing state, before acclimatization, and when the rectal temperature is elevated, significant increases in resting metabolism occur and that these increases are in general in accord with the Van't Hoff coefficient (106). After acclimatization, however, low levels are a more uniform finding (107). A detailed account of the course and magnitude of the lowering in resting metabolism has been reported (108). In contrast to the above, others (21, 109) found no change in basal metabolism after from 5 to 10 days' work in the heat; the subjects, however, exhibited elevations in rectal temperature at the time.

A. *Caloric requirements and specific dynamic action of proteins.* No great differences in the caloric requirement of man occur as the result of high temperatures. Studies on rats have indicated a lessened requirement at high tempera-

ture as evidenced by greater weight gain with similar intake and activity, and study of the voluntary food intake of United States soldiers did not reveal any measurable seasonal differences (110). This is in harmony with the observations that no significant changes in efficiency of working metabolism as a result of heat are demonstrable at high rates of caloric expenditure. Any differences that might be expected to result from the lessened requirement for heat production per se, would be quantitatively insignificant in relation to the overall expenditure. Similarly, any heat production from the specific dynamic action of proteins is of no consequence. That it moreover is not measurable has been shown (109).

B. *Vitamin requirements.* The situation as to optimal intake of vitamins and the benefits of supplements has been greatly clarified. Mills and colleagues (111, 112, 113, 114, 155) believe that there is increased requirement in the heat with increased need for choline (5 x normal). It is doubted that these results are valid as applied to man. Well-controlled human studies by others (109, 115, 116, 117, 118, 119, 120, 121) offer convincing evidence that there is no increase in requirement or benefit from added supplements of vitamins or choline.

Whether inferences as to requirement may be drawn from excretion levels alone is doubtful. Dietary intake is but one facet of a figure which comprises need, intake, absorption, distribution, utilization, excretion, etc. With clinical observations valid for at least the duration of the studies, one is led to the conclusion that isolated observations on animals and on excretion, etc., are interesting but not definitive with respect to man.

The possibility of increasing the ability of man to work in the heat by means of administration of specific substances has been attractive. The results are disappointing (122, 123). Many of the symptoms of heat exhaustion are like those of adrenal insufficiency (124). This, and other reports (125) on the beneficial effects of adrenal cortex in hyperthermia led to its trial in men working in hot environments. No consistent effects were noted (97, 105).

7. *Gastro-Enteric Tract.* The anorexia, nausea, vomiting, cramps and frequent diarrhea that often occur during high heat stress, reveal the pronounced effects which heat may have upon gastro-enteric tract function. Explanation for the occurrence of these phenomena is not complete. The vomiting has been explained on the basis of rejection by the stomach of fluid that is not being absorbed because of limitation in blood flow to the stomach (88). Henschel et al. (24) studied the gastric emptying time of 17 normal men at temperatures of from 77 to 120°F. In all but 1 the gastric emptying time was faster at the higher temperature; in 12 of the 17 a decrease of 30 per cent in time was seen at the higher temperatures. A study of 100 men doing hard work at 120°F. failed to reveal any anorexia or decreased gastric activity except when actual heat exhaustion occurred.

8. *Thermal Regulation.* Thermal regulation continues to be a matter of great interest and has recently been reviewed (3, 46, 126, 156). Further light upon the factor of chemical regulation is offered in the report of Hardy (15). Other reports (46, 127) have been concerned with the character of the nervous control and thyroid activity. A significant paper covering clinico-pathologic studies of

disturbances of temperature regulation in man has been presented by Davison (128). Four cases of hyperthermia and one case of hypothermia were studied. In three cases of hyperthermia the lesions responsible were sharply circumscribed in the ventromedial tuberal nuclei in the floor of the 3rd ventricle.

9. *Acclimatization.* Man is a homothermic animal capable of great increases in internal heat production. He, moreover, lives in a constantly changing thermal environment, all of which requires continuous regulation to maintain thermal equilibrium at his normal level of internal temperature. Even short bursts of work with its high heat production or minutes of interference with heat dissipation (hot rooms, excess clothing, etc.) will increase heat storage and elevate internal temperature (154). With sustained increased level of heat production or continued interference with dissipation, a new level of thermal equilibruium is established by the individual. As this is repeated or continued for days there is rapid acquisition of the ability to maintain the same or lower level of thermal equilibrium under the same conditions of stress, at less cost to the individual. When the shift in temperature is moderate, the initial load on the homeostatic mechanisms is minimal and often unnoticed. Slight increases in rectal temperature and pulse rate are nonetheless usual (129). When the change is great, and to an environment of high thermal stress, with initial marked physiological effects, the subsequent improvement in performance and reduction in load upon homeostatic mechanisms is spectacular. This acquired capacity entails complex readjustments, the mechanism for which is often conjectural. It has been suggested that shift of vasomotor control to thalamic centers is entailed (22). Without regard to hypothesis, the process of acclimatization is characterized by certain phenomena and changes which have been well studied and described in investigations carried out within tolerable thermal limits (21, 23, 24, 25, 33, 41, 130, 132, 133).

A. *General state.* Beginning with the first day and continuing at a decreasing rate for from 8 to 10 days there is progressive improvement in the appearance and behavior of men. Incidence of heat exhaustion, headache, dizziness, polypnea, gastroenteric tract symptoms, inco-ordination, irritability and depression decrease progressively. Most of the improvement occurs in the first 4 or 5 days and when acclimatization is achieved the subject will often perform as willingly and well in the heat as was formerly possible under temperate conditions.

B. *Circulation.* Circulatory system instability which is closely related to the production of many of the above symptoms, follows the same general course during the acclimatization (21, 33, 41, 71, 72, 131). Changes in heart rate are dramatic, the principal reduction occurring in the first 4 or 5 days.

C. *Rectal temperature.* The general course of the rectal temperature is progressively downward during acclimatization, the reduction being more gradual than that of changes in circulation or performance. At low levels of heat stress, normal levels may be reached in from 7 to 10 days—in general, levels of from a few tenths to a degree or more above normal still are encountered after 10 days.

D. *Sweating.* Rates of sweating uniformly increase during acclimatization

though the magnitude of increase is widely variable. Increases in rates of sweating usually begin later (after from 3 to 5 days) than do other changes and may continue longer and even augment later (41, 52). The lack of relation between this phenomenon and the mean skin and rectal temperatures was demonstrated in Horvath's (52) study of 16 men acclimatized to severe conditions (T_w93°F.). Chloride concentration tends to decrease during acclimatization (134) except in the presence of exhausting work.

There has been good agreement among most observers that the performance and appearance of men, degree of vascular instability, cardiac rate and rectal temperatures when evaluated together, are so closely related to the state of acclimatization as to comprise the most reliable indices of the state.

The practical limits of work and environment for which acclimatization can be achieved are reviewed below. Limits for acclimatization are not only sharply defined but the adjustments involved in the acclimatization are easily disturbed. Thus, a fully acclimatized man may exhibit any or all of the disturbances characteristic of the unacclimatized state if excessive rates of work are imposed (21, 23, 24, 33, 71, 135) or if there is intercurrent infection (7), loss of sleep, or alcoholic indulgence (21, 33). More importantly, failure to maintain complete hydration (18, 21, 25, 136), caloric intake (120) or salt intake (11, 41, 104) will lead to prompt deterioration in performance. The same factors which enable good performance in the cool and which are related to good initial capacity to perform in, and acclimatize to heat, are similarly important to the maintenance of the acclimatization. The acclimatized state is not only unstable with respect to the above factors but is also a temporary adjustment in that it persists for a limited time only. With reduction in thermal stress, acclimatization is lost more slowly than it was acquired, the rate being influenced directly by the magnitude of the difference in the ambient temperatures (21, 24). A good state of acclimatization is maintained for 1 or 2 weeks with gradual loss thereafter; some men retain a fair degree at the end of 2 months. In studies in the winter months persistence for at least 3 weeks is usual (71). Repeated exposures at least 1 month apart are required to maintain good acclimatization (21).

E. *Factors influencing acclimatization.* A number of factors influence the ability to acquire acclimatization and the rate of its development. Resting in the heat is associated with a limited adjustment, but work within limits of tolerance is necessary for full development. Exposure and work in hot, dry environments results in a partial acclimatization to hot humid conditions. Repeated brief (1½ to 3 hrs.) daily bouts of work in the heat will acclimatize men but fullest adjustment is attained most quickly by graded, progressively increasing work in the heat.

After acclimatization to one level of thermal stress, further acclimatization may be acquired for more severe environments (23, 33), and acclimatization to a high level of heat stress enables greatly enhanced performance in less severe heat (52).

Long-term changes have not been studied under controlled conditions though the performance of men after months or years of work and residence in environ-

ments of high thermal stress is a matter of great practical importance and short-term studies cannot be expected to elucidate these slowly changing factors. These may be associated with deterioration, rather than improvement in adjustment and performance (4,137,138). Several reports (97,98,99) have called attention to the instability of the sweating mechanism that may appear after months of residence in the heat and Collings et al. (96) noted stoppage of sweating in 26 per cent of steelworkers prior to breakdown—many of the men having had years of exposure. Certainly when the stress is high, deterioration can occur in a matter of hours or days (151, 152, 153). The character of the changes at low orders of stress maintained for years is not indicated. Investigation of the problem of deterioration is most urgent.

10. *Work in Hot Environments.* Work in hot environments imposes two additive stresses, increased internal heat production, and added demands upon the circulation for the supply of muscle. Determinations of the relationship between the stresses of ambient conditions, work and dehydration have been reported and reviewed (19, 133, 137, 139, 140).

11. *Upper Limits of Tolerance.* All of the factors which influence environmental stress, heat dissipation, heat production or the fitness of the man, become qualifying variables in any attempt to define limits of adjustment. In addition, the criterion of limit may range from discomfort to death. The relative significance of certain of these factors has been studied and referred to above: Environmental; air temperature, water vapor in air, duration of exposure, air movement, radiation, clothing, constancy of stress. Individual; diseases, work, body size, acclimatization, age, fitness, individual variation, posture, fatigue, water balance, salt intake, caloric intake, alcohol. It is apparent that the large number of variables will limit the application of results of particular experiments rather narrowly to the conditions and types of subjects employed. A further complication arises in definition of limit of tolerance. Should some level of performance be employed or should the convenient indices of pulse and rectal temperture be used? If the latter be employed and we accept the rectal temperature during the steady state as the index of the existing level of thermal equilibrium, one faces the question of what constitutes an acceptable range of increase in levels. Certain limiting values have been proposed (23, 133, 140). Attempts to relate psychophysiological performance to ambient conditions and to level of rectal temperature have not yielded very promising results (141).

Studies of Eichna et al. (33) have established a series of environmental conditions at which sustained work (4 hrs., at 250 cal./hour) is easily possible, difficult or impossible as a day-to-day performance for periods up to one week. Their subjects were well trained, fully acclimatized, young men, and were maintained on adequate salt and water intakes. They found, as had Haldane, that the T_w was the most important factor in determining the upper limits of tolerance and a better index of climatic stress up to 120°T_d than was T_d or the effective temperature. At T_w up to 91°*F*, men worked easily and efficiently, maintaining thermal equilibrium with rectal temperatures under 101°F. Between this tolerable ambient level and T_w 94°F (above which sustained work was impossible),

work became progressively more difficult and heat casualties occurred with increasing frequency. There is general agreement as to the level of tolerable T_w. In 1909 Haldane (142) set the upper limit of tolerance for resting men in still air at T_w 90°F as did McConnell, Houghton and Yaglou in 1924 (35). Other reports (30, 38, 143, 144) arrived at similar limits. The agreement in level is in general fairly good, but not perfect. Robinson (23) reported a series of 212 experiments carried out at temperatures ranging from 23 to 50°C with various humidities. His subjects maintained work at metabolic rates of 130 cal. M²/hour at 35°C with 96 per cent relative humidity and were in thermal equilibrium from the second to sixth hours of exposures at 34°C with 91 per cent relative humidity. The ability to maintain thermal equilibrium was taken as the criterion for tolerance. We believe that the maintenance of equilibrium is a good measure of the capacities of the homeostatic mechanisms but is not itself a measure of limits of tolerance since new thermal equilibria may be established at any level within the ability of the individual to withstand the thermal stress.

Search has continued for means of evaluation of total thermal stress in terms of known ambient conditions and the correlation of values obtained with physiological effects. Robinson (23) arrived at an index of physiological effect of environment from weighted data on heart rate, skin temperature, rectal temperature and rate of sweating. This he designated as E_p or index of physiological effect. Using this index, environmental conditions which resulted in equal E_p values were determined and are presented in graphic form. Differences in behavior of clothed and nude men are given. With severe thermal stress Robinson's contours parallel those of Houghton and Yaglou (145) for effective temperatures. In comparison with the results of Gagge, Harrington, and Winslow (30), Robinson's contour line extrapolated to 100 per cent relative humidity, is about one degree higher at saturation with about the same air movement.

Plummer, Cochrane and Siple have developed a similar concept, designated as "Thermal Acceptance Ratio" (54). Methods of evaluating environmental loads are presented. Like other approaches, this requires the assumption of a skin temperature and makes no provision for the complex factors involved in sweating and evaporation. Since the stress imposed by an environment is a function of the ratio between the ability of the environment to take up heat and the requirement for heat output by the body, considerably more information is needed before refined general expressions will be entirely acceptable. A considerable advance in defining relationships has been made by Nelson and coworkers (28).

Much-needed observations on the effects of high levels of radiation gain have been reported (135, 146, 148, 149). The increases in metabolic rate caused by radiation are significant.

12. *Heat Disease.* In heat stress, either by reason of the direct effect of high levels of thermal equilibrium or because of overstressing of homeostatic mechanisms, performance may quickly deteriorate and collapse occur. Again, with high stress, there may be a gradual deterioration in work capacity with the appearance of distressing and incapacitating symptoms. With this breakdown,

and without regard to type of onset, there appears a train of physiological disturbances which, when taken together, make up the clinical syndromes of heat cramps, heat exhaustion, and heat stroke. The syndromes are descriptive of the final states encountered and are practically useful in determining therapy; physiologically their occurrence throws light upon the character of the breakdown of the overstressed homeostatic mechanisms. The conditions encountered clinically have been described in detail (11, 36, 37, 39, 40). The most spectacular of these syndromes, heat stroke, is characterized by coma, hyperpyrexia, cessation of sweating and circulatory failure. Levels of hydration and salt content are normal and extreme levels of rectal temperature (108–112°F.) are encountered. The second type, commonly differentiated into 2 syndromes, heat cramps and heat exhaustion, appears to have a common basis, the pattern of symptoms being determined by the relative importance of the effects caused by circulatory system instability, salt deprivation, and dehydration (11, 41, 96). The importance of cessation of sweating has been pointed out above—it is frequently associated with impending failure or stroke and may cause collapse with or without hyperthermia. The question of fatigability of the mechanism has been investigated; the experiences in the field suggest this is an important consideration and the reduction in capacity with time or excessive load of the homeostatic mechanisms is demonstrated in the studies of environments of high thermal stress (23, 24, 33, 151, 152, 153).

Local metabolic and functional abnormalities may result from altered hemodynamics, deficient blood supply, or from the direct effect of temperature itself. Additional evidence of the damaging effects of heat upon the brain have been reported (147). Oxygen uptake of rat cerebral cortex was found to be greatly reduced at temperatures of from 40–47.5°C (104–116°F.). The rate of utilization decreased more rapidly with time as the temperatures were increased, becoming asymptotic at a level 10 per cent of that of control slices at 37.5°C. In about half the experiments the oxygen consumption was constant for an hour or more, in the remainder there was decrease with time beginning as early as 30 minutes after thermal equilibrium. The inhibition was reversible after 3 hours' exposure at 40°C. (104°F.) but after even 30 minutes' exposure at 40.8°C. (105.4°F.) or higher, recovery was incomplete after cooling. Temperatures higher than 41°C. resulted in progressively less complete recovery. These findings are consistent with reports of sequelae following hyperthermia. These have been reviewed (150).

SUMMARY

No conclusions are drawn but the general pattern of the investigations prompts certain comments.

The researches of the last few years have demonstrated that the performance and state of man in relation to his environment is a measurable phenomenon with qualitative aspects often quite dissimilar from the measurable differences in the components that together make up the whole. Man as a functioning organism is quite different in character than any of the systems into which, in an

analytic approach, he may be resolved. The lesson for physiologists is clear; the overall performance of the integrated being should be the point of departure for research endeavors, and the yardstick by which the individual results are gauged. Both types of activity are essential—it is only by splitting off bits that it becomes possible to provide the bricks for building, and only by continuing and related study of the overall performance do we provide the observations necessary for definition of the individual problems, and the architecture of the whole.

REFERENCES

(1) Bazett, H. C. Physiol. Rev. 7: 531, 1927.
(2) Mom, E. P. Arch. f. Schiffs. u. Inf. Hyg. 44: 182, 1940.
(3) Adolph, E. F. Physiological Regulations. Jacques Cattell Press, Lancaster, Pa., 1943.
(4) Sundstroem, E. S. Physiol. Rev. 7: 320, 1927.
(5) Chambers, A. L. Cand. M. A. J. 48: 214, 1943.
(6) Machle, W. Mil. Surg. 95: 98, 1944.
(7) Morton, T. C. Trans Roy. Soc. Trop. Med. Hyg. 37: 317, 1944.
(8) Maclean, K. S. J. Roy. Nav. Med. Serv. 29: 31, 1943.
(9) Critchley, M. Brit. M. J. 2: 173, 1945.
(10) Eichna, L. W., R. H. Walpole, W. B. Shelley and J. L. Whittenberger. Armored Med. Res. Lab., Fort Knox, Report on Project #35, (Sept. 13), 1944.
(11) Stenning, J. C. J. Roy. Nav. Med. Serv. 31: 120, 1945.
(12) Ashve. Report to the Bureau of Ships U.S.N., (July) 1941.
(13) Memo. Report of Aero. Med. Lab. Air Tech. Serv. Command—Wright Field, 3-695-49A, 1945.
(14) Hardy, J. D. and E. F. DuBois. Proc. Nat. Acad. Sc. 26: 389, 1940.
(15) Hardy, J. D., A. T. Milhorat and E. F. DuBois. J. Nutrition 21: 383, 1941.
(16) Day, R. Am. J. Dis. Child. 61: 734, 1941.
(17) Burch, G. E. Am. J. Med. Sc. 211: 181, (Feb.), 1946.
(18) Adolph, E. F. Fed. Proc. 2: 158, 1943.
(19) Adolph, E. F. Physiology of Man in the Desert. Interscience Pub., in press.
(20) Johnson, R. E. and S. Robinson. Contract cmr.—Report 16, (15 Feb.), 1946.
(21) Bean, W. B. and L. W. Eichna. Fed. Proc. 2: 144, 1943.
(22) Johnson, R. E. Ann. Rev. Physiol. 8: 535, 1946.
(23) Robinson, S., E. S. Turrell and S. D. Gerking. Am. J. Physiol. 143: 21, 1945.
(24) Henschel, A., H. L. Taylor and A. Keys. Am. J. Physiol. 140: 321, 1943.
(25) Eichna, L. W., W. B. Bean, W. F. Ashe and N. Nelson. Bull. Johns Hopkins Hosp. 76: 25, 1945.
(26) Winslow, C-E.A., L. P. Herrington and A. P. Gagge. Am. J. Physiol. 116: 669, 1936.
(27) Winslow, C-E.A., A. P. Gagge and L. P. Herrington. Am. J. Physiol. 127: 505, 1939.
(28) Nelson, N., L. W. Eichna, H. M. Horvath, W. B. Shelley and T. F. Hatch. J. Clin. Investigation, in press.
(29) Gagge, A. P. Am. J. Physiol. 120: 277, 1937.
(30) Gagge, A. P., L. P. Herrington C-E.A. Winslow. Am. J. Hyg. 26: 84, 1937.
(31) Burton, A. C. J. Nutrition 9: 261, 1935.
(32) Hardy, J. D. and E. R. DuBois. J. Nutrition 15: 477, 1938.
(33) Eichna, L. W., W. F. Ashe, W. B. Bean and W. B. Shelly. J. Ind. Hyg. Toxicol. 27: 59, 1945.
(34) Winslow, C-E.A., L. P. Herrington and A. P. Gagge. Am. J. Physiol. 120: 1, 1937.

(35) McConnell, W. J., F. C. Houghten and C. P. Yaglou. J. Am. Soc. Heat. Vent. Eng. **30**: 167, 1924.
(36) Bottner, A. Klin. Wchnschr. **20**: 471, 1941.
(37) Wallace, A. W. Mil. Surgeon **93**: 140, 1943.
(38) Caplan, A. Bull. 54, Kolar Gold Field Min. Met. Soc.
(39) Croom, J. H. J. Roy. Army Med. Corps **83**: 288, 1944.
(40) Ladell, W.S.S., J. C. Wattenlove and M. F. Hudson Lancet **2**: 491, 1944.
(41) Taylor, H. L., A. Henschel, O. Mickelson and A. Keyes. Am. J. Physiol. **140**: 439, 1943.
(42) Haldane, J. S. J. Hyg. **5**: 495, 1905.
(43) Dreosti, A. O. J. Chem. Met. Min. Soc. South Africa **36**: 102, 1935.
(44) Bernard, C. Lecons sur la Chaleur Animal. Paris, Bailliere, 1876.
(45) Bazett, H. C. and B. McGlone. Am. J. Physiol. **82**: 415, 1927.
(46) Herrington, L. P. and A. P. Gagge. Ann. Rev. Physiol. **5**: 295, 1943.
(47) Spealman, C. R. Proc. Soc. Exper. Biol. Med. **60**: 11, 1943.
(48) Manshak, M. E. Arch Biolog. Nauk. (Leningrad) **55**: 20, 1939.
(49) Fedorov, N. A. and E. D. Shur. Am. J. Physiol. **137**: 30, 1942.
(50) Milhorat, A. T., J. D. Hardy and A. J. Fair. J. Clin. Investigation **21**: 642, 1942.
(51) Barcroft, H. and O. G. Edholm. J. Physiol. **102**: 5, 1943.
(52) Horvath, S. M. and W. B. Shelley. Am. J. Physiol. **146**: 336, (June) 1946.
(53) Carley, C. L. and J. L. Nickerson. Am. J. Physiol. **143**: 373, 1945.
(54) Plummer, J. H., M. I. Cochrane and P. A. Siple. Climatol. Environ. Protect. Sect., O.Q.M.G., (Aug. 20), 1945.
(55) Hardy, J. D. and E. F. Dubois. J. Nutrition **15**: 461, 1938.
(56) Burton, A. C. and H. C. Bazett. Am. J. Physiol. **117**: 36, 1936.
(57) Winslow, C-E.A., L. P. Herrington and A. P. Gagge. Am. J. Physiol. **120**: 1, 1937.
(58) Houghten, F. C. and C. P. Yaglou. ASHVE, Trans. **29**: 361, 1923.
(59) Roth, G. M., B. T. Horton and C. Sheard. Proc. Staff, Meet. Mayo Clin. **15**: 774, 1940.
(60) Sheard, C. and M. M. Williams. Proc. Staff, Meet Mayo Clin. **15**: 758, 1940.
(61) Kirklin, O. L., W A. Plummer and C. Sheard. Am. J. Physiol. **128**: 782, 1940.
(62) Horton, B. T., C. Sheard and G. M. Roth. J. Med. Soc. New Jersey **37**: 311, 1940.
(63) Hyndman, O. R. and J. Wolkin. Am. Heart J. **22**: 280, 1941.
(64) Naide, M. Am. J. Med. Sc. **207**: 606, 1944.
(65) Forster, R. E., B. G. Ferris and R. Day. Am. J. Physiol. **146**: 600, (July) 1946.
(66) Abramson, I., Vascular Responses in the Extremities of Man in Health and Disease. University Chicago Press, Chicago, Ill., 1944.
(67) Spealman, C. R. Am. J. Physiol. **145**: 218, 1945.
(68) Stewart, H. J. and W. F. Evans. Am. Heart J. **26**: 67, 1943.
(69) Freedberg, A. S., E. D. Spiegl and J.E.F. Riseman. Am. Heart J. **27**: 611, 1944.
(70) Proskauer, G. G., C. Neuman and A. de Graff. Am. J. Physiol. **143**: 290, 1945.
(71) Taylor, H. L., A. Henschel and A. Keys. Am. J. Physiol. **139**: 583, 1943.
(72) Eichna, L. W. In preparation.
(73) Hellebrandt, F. A. and E. B. Franseen. Physiol. Rev. **23**: 220, 1943.
(74) Maclean, A. R., E. V. Allen and T. B. Magath. Am. Heart J. **27**: 145, 1944.
(75) Starr, I. J. Clin. Investigation **22**: 813, 1943.
(76) Allen, S., H. L. Taylor and V. C. Hall. Am. J. Physiol. **143**: 11, 1945.
(77) Wiggers, H. C., G. H. Glasser, K. De S. Canavarro and A. E. Treat, Am. J. Physiol. **139**: 217, 1943.
(78) Warren, J. V. and E. A. Stead. Am. J. Med. Sc. **205**: 501, 1943.
(79) Henderson, Y. Medicine **22**: 223, 1943.
(80) Blalock, A. and M. F. Mason. Arch. Surg. **42**: 1054, 1941.
(81) Antes, R. J. Proc. Soc. Exper. Biol. Med. **56**: 60, 1944.

(82) Elman, R., W. M. Cox, Jr., C. E. Lischer and A. J. Mueller. Proc. Soc. Exper. Biol. Med. 51: 350, 1942.
(83) Steele, J. Internat. Clinics 1: 17, 1936.
(84) Steele, J. and A. E. Cohn. J. Clin. Investigation 13: 869, 1934.
(85) Kinsey, D. and P. D. White. Arch Int. Med. 65: 163, 1940.
(86) Krogh, A. Scand. Arch. f. Physiol. 27: 227, 1912.
(87) Wilkins, R. W., J. S. Hunt and C K. Friedland. J. Clin. Investigation 21: 625, 1942.
(88) Amussen, E. Am. J. Physiol. 131: 54, 1940.
(89) Glickman, N., F. K. Hick, R. W. Keeton and M. M. Montgomery. Am. J. Physiol. 134: 165, 1941.
(90) Newburgh, L. H. and M. W. Johnson. Physiol. Rev. 22: 1, 1942.
(91) Barbour, H. G. Res. Pub. Assoc. Res. Nerv. Ment. Dis. 20: 449, 1940.
(92) Pinson, E. A. Am. J. Physiol. 137: 492, 1942.
(93) Burch, G. E. and W. A. Sodeman. Am. J. Physiol. 138: 603, 1943.
(94) Beaton, L. E., W. A. McKinley, C.M. Berry and S. W. Ranson. J. Neurophysiol. 4: 478, 1941.
(95) Gurney, R. and I. L. Bunnel. J. Clin. Investigation 21: 269, 1942.
(96) Collings, G. H., L. A. Shoudy and F. E. Shaffer. Indust. Med. 12: 728, 1943.
(97) Ladell, W. S. S. J. Physiol. 104: 13P, 1945.
(98) Novy, F. G. and J. H. Ramsey. J.A.M.A. 125: 738, 1944.
(99) Wolkin, J., J. I. Goodman and W. E. Kelley. J.A.M.A. 124: 478, 1944.
(100) Sunderman, F. W. Arch. Int. Med. 67: 846, 1941.
(101) Sodeman, W. A. and G. E. Burch. J. Clin. Investigation 23: 37, 1944.
(102) Talbott, J. H. Medicine 14: 323, 1935.
(103) Report of Council of Pharmacy & Chemistry, J.A.M.A. 129: 131, 1945.
(104) Pitts, G. C., R. E. Johnson and F. C. Consolazio. Am. J. Physiol. 142: 253, 1944.
(105) Moreira, M., R. E. Johnson, A. P. Forbes and F. C. Consolazio. Am. J. Physiol. 143: 169, 1945.
(106) DuBois, E. F. J.A.M.A. 77: 352, 1921.
(107) DuBois, E. F. Bull. N.Y. Acad. Med. 15: 143, 1939.
(108) Mason, E. D. Indian J. Med. Res. 32: 27, 1944.
(109) Pitts, G. C., F. C. Consolazio and R. E. Johnson. J. Nutrition 27: 497, 1944.
(110) Howe, P. E. and G. H. Berryman. Am. J. Physiol. 144: 588, 1945.
(111) Mills, C. A. Am. J. Physiol. 133: 525, 1941.
(112) Mills, C. A. Arch Biochem 1: 73, 1942.
(113) Mills, C. A. Am. J. Trop. Med. 25: 59, 1945.
(114) Mills, C. A. Arch. Biochem. 3: 333, 1944.
(115) Henschel, H. F., H. L. Taylor and A. Keys. Am. J. Trop. Med. 24: 259, 1944.
(116) Shoudy, L. A. and G. H. Collings. Ind. Med. 14: 573, 1945.
(117) Keys, A. Fed. Proc. 2: 164, 1943.
(118) Holt, L. E. Southern Med. Surg. 109: 9, 1943.
(119) Osborne, S. L. and C. J. Farmer. Proc. Soc. Exper. Biol. Med. 49: 575, 1942.
(120) Johnson, R. E. Gastroenterol 1: 832, 1943.
(121) Sargent, F., P. Robinson and R. E. Johnson. J. Biol. Chem. 153: 285, 1944.
(122) Okusa, H. J. Sci. Labour (Kurashiki) 16: 867, 1930.
(123) Wills, J. H. Fed. Proc. 4: 77, 1945.
(124) Edelman, A., D. H. Mahanna, L. A. Lewis, J. S. Thatcher and F. A. Hartmann. J. Clin. Endocrinology 3: 20, 1943.
(125) Schlegel, B. Klin. Wchnschr. 20: 506, 1941.
(126) Herrington, L. P. Am. J. Physiol. 129: 123, 1940.
(127) Lelcaloh, A. B. Bull. Biol. Med. Exp., USSR. 8: 194, 1939.
(128) Davison, C. Assoc. Res. Nerv. Ment. Dis. Proc. 20: 774, 1940.
(129) Behnke, A. R. Mil. Surgeon 88: 633, 1941.

(130) ADOLPH, E. F., H. H. BROWN AND J. TOWBIN. Interim Report #17 to Comm. on Med. Res., (June 1), 1944.
(131) ROBINSON, S., E. S. TURRELL, H. S. BELDING AND S. M. HORVATH. Am. J. Physiol. 140: 168, 1943.
(132) SHELLEY, W. B., L. W. EICHNA AND S. M. HORVATH. J. Clin. Investigation 25: 437 (May), 1946.
(133) FLEISHER, W. L., A. E. STACEY, F. C. HOUGHTEN, AND M. B. FERDERBER. Heat, Pip. and Air Cond., (Feb., Mar., Apr.), 1939.
(134) JOHNSON, R. E., G. C. PITTS, AND F. C. CONSOLAZIO. Am. J. Physiol. 141: 575, 1944.
(135) GOSSELIN, R. E. Fed. Proc. 4: 25, 1945.
(136) ADOLPH, E. F. Fed. Proc. 4: 1, 1945.
(137) EIJKMAN, C. Lancet 1: 887, 1924.
(138) Editorial, Bull. U.S. Army Med. Dept. 82: 10, 1944.
(139) BRUNT, D. Quat. J. Roy. Meteor. Soc. 69: 77, 1943.
(140) HOUGHTEN, F. C., M. B. FERDERBER AND A. A. ROSENBERG. Ind. Med., Ind. Hyg. Section 1: 7, 1940.
(141) WEINER, J. S. AND J. C. D. HUTCHINSON. Brit. J. Ind. Med. 2: 154, 1945.
(142) HALDANE, J. S. British Committee on Humidity and Saturation in Cotton Weaving Sheds, London, 1909.
(143) WINSLOW, C-E.A., L. P. HERRINGTON, AND A. P. GAGGE. Am. J. Physiol. 120: 288, 1937; 124: 692, 1938.
(144) KLUGE, H. Arbeitsphysiol. 11: 101, 1940.
(145) HOUGHTEN, F. C. AND C. P. YAGLOU. J. Am. Soc. Heat. Vent. Eng. 29: 165, 1923.
(146) HILL, L. Brit. J. Physiol. Med. 5: 128, 1942.
(147) FIELD, J., F. A. FUHRMAN AND A. W. MARTIN. J. Neurophysiol. 7: 117, 1944.
(148) LIPSON, N. AND M. B. VISSCHER. J. Ind. Hyg. Toxicol. 25: 484, 1943.
(149) BLUM, H. F. Physiol. Rev. 25: 483, 1945.
(150) FREEMAN, W. AND S. E. DUMOFF. Arch. Neurol. Psych. 51: 67, 1944.
(151) ROBINSON, S. AND S. D. GERKING. Am. J. Physiol. In press.
(152) ROBINSON, S., S. D. GERKING AND E. S. TURRELL. Am. J. Physiol., in press.
(153) GERKING, S. D. AND S. ROBINSON. Am. J. Physiol., in press.
(154) NIELSEN, M. Skand. Arch. Physiol. 79: 193, 1938.
(155) MILLS, C. A., E. COTTINGHAM AND M. MILLS. Am. J. Physiol. 141: 359, 1944.
(156) FUHRMAN, F. A. Physiol. Rev. 26: 247, (Apr.) 1946.
(157) BURCH, G. E. Proc. Cent. Soc. Clin. Res. 18: 60, 1945.
(158) LADELL, W.S.S. Brit. Med. Bull. 3: 175, 1945.

THE DEFINITION AND MEANING OF pH

A. G. OGSTON

Department of Biochemistry, Oxford, England

This is intended to be a critical review of the meanings that are attached to the term "pH". Some of the principles underlying the methods of measuring it are discussed, but not their experimental details.

There is need for a review of this sort since "pH" may bear several different meanings and since, in spite of its universal use, it is clear that many of its users are not aware of this fact. It may be an exact or inexact statement of the result of a given method of measurement; it may represent an exact physico-chemical quantity; often it is used in senses which are physicochemically inexact and ambiguous.

The object of this review is to examine the possible meanings of "pH" and their respective usefulness in the light of well established physicochemical principles. A convention of expression is suggested which should promote clarity and prevent confusion.

I. *Bulk Phases.* pH is most commonly used as a term of description for ordinary bulk solutions. It may serve two rather different purposes: first, to define empirically a unique property of a solvent mixture (including buffer and neutral salts etc.) which shall determine the state of dissociation equilibrium of any given acid dissolved in it, in such a way that the same state of dissociation can be reproduced at will: secondly, to form part of an exact physico-chemical expression of the state of dissociation.

The importance of the hydrogen ion as a factor controlling the dissociation of acids and bases in solution was recognised when Arrhenius applied the law of Mass Action to ionic equilibria. Later it was realised that many substances of biological importance are weak acids or bases whose degrees of dissociation in part determine their biological activity.

In the modern terminology of Bronsted (1) and Lowry (2), any acid-base equilibrium may be expressed as

$$A \rightleftharpoons B + H^+$$

where A and B are conjugate acid and base and H^+ represents, in aqueous solution, the ion H_3O^+. Application of the classical Mass Action theory gives

$$K = (H^+)\cdot(B)/(A) \tag{1}$$

Sørensen (3) introduced the term pH as a convenient shorthand for $-\log(H^+)$; equation (1) is transformed to

$$\mathrm{pH} = pK + \log\,(B)/(A) \tag{2}$$

The use in this sense of pH was typographically convenient; it allowed the whole range of solutions, from strongly acid to strongly alkaline, to be expressed on a compact graphical scale; equation (2) gives the useful symmetrical dissociation,

ionisation or titration curve and furthermore shows that pH defines the value of $(B)/(A)$ for a given value of K.

However, it became evident that equation (1) is not a generally valid expression of acid-base equilibria; the values of K were found to vary with concentrations of the acid-base system and of neutral salts. G. N. Lewis (4) therefore introduced the concepts of "activity", a, and "activity coefficient", f. In ionic solutions these are usually related by

$$a = f.c$$

where c is the concentration; f is defined as approaching unity at infinite dilution and at finite concentration is given such values as to make the value of K equal to that expressed in terms of concentration at infinite dilution. Equation (1) becomes

$$K = \frac{a_{H^+} \cdot a_B}{a_A} \text{ or } \quad K = (H^+) \cdot (B)/(A) \times \frac{f_{H^+} \cdot f_B}{f_A} \tag{3}$$

We shall use the second form in the further discussion.

The logarithmic form of (3) analogous to (2) is

$$\log K = \log (H^+) + \log (B)/(A) + \log \frac{f_{H^+} \cdot f_B}{f_A} \tag{4}$$

It contains four terms in place of the three of equation (2). The "p" notation has purposely not been used because there are two questions which first require answers:

1. Can equation (4) be satisfactorily reduced to three terms, by incorporation of the activity coefficients in one or more of the other terms, so that it becomes like equation (2)?
2. What are the quantities in equation (4) that are measured by the means used for measuring "pH"?

These questions should be considered together because it is desirable that the result of an experimental determination of "pH" should correspond with one of the terms of equation (4) or of a modification of it.

Consider first the ways in which equation (4) might be reduced to three terms.

1. The activity coefficients might be distributed amongst the other three terms, but this may be done only in certain ways. Guggenheim (5) has shown conclusively that the activity or activity coefficient of a single sort of ion cannot be unambiguously measured or defined; this applies more generally to any function

$$\frac{(f_{x})^{m} (f_{y})^{n} (f_{z})^{o} \cdots}{(f_{x'})^{m'} \cdot (f_{y'})^{n'} \cdot (f_{z'})^{o'} \cdots} \tag{5}$$

such that the algebraic sum of the ionic charges of $x, y, z \cdots$, each multiplied by its exponent *minus* that of the charges of $x', y', z' \cdots$ is not zero. The attachment of any numerical value to such a function must involve some arbitrary assumption; and while the use of such a function may sometimes be con-

venient, it is easy to overlook the fact that an implicit assumption has been made and, consequently, to become involved in an argument which can lead only to that assumption, perhaps in a disguised form.

The charge of B is always one unit more negative than that of A. It follows that there are only two cases in which the division of the activity coefficients is legitimate; that in which A carries no charge, when it is permissible to separate $f_{H^+} \cdot f_{B^-}$ from f_A, leading to

$$pH = pK + \log \frac{(B^-)}{(A) \cdot f_A} \tag{6}$$

where pH $= -\log(H^+) \cdot f_{H^+} \cdot f_{B^-}$;
and that in which B carries no charge, which leads to

$$pH = pK + \log \frac{(B) \cdot f_B}{(A^+)}$$

where pH $= -\log(H^+) \cdot f_{H^+}/f_{A^+}$. (7)

These two definitions of pH correspond with the quantities that are obtained in certain measurements and will be discussed further.

One definition of pH is often used which is quite illegitimate; this is

$$\text{pH} = -\log(H^+) \cdot f_{H^+} \tag{8}$$

Authors frequently refer to the "activity of hydrogen ion" or the "activity coefficient of hydrogen ion" apparently thinking that they are being peculiarly precise and unaware that they are using terms which are physically undefinable and whose magnitudes are experimentally unmeasurable, unless some special arbitrary assumption is made about the magnitude of some other quantity in the system.

One such assumption may perhaps be justifiable, namely that the properties of any ion depend only on its charge. This may be approximately true in the range of concentrations to which the Debye-Huckel treatment applies, though even in such dilute solutions the radii of the ions have some effect. The mean activity coefficient of any uni-univalent electrolyte may then be defined as

$$f_+ \cdot f_- = f_\pm^2$$

and

$$\text{pH} = -\log(H^+) \cdot f_\pm$$

This is sometimes a useful convention; but it should be recognised as being arbitrary.

2. The activity coefficients might be combined with K to give a modified, variable "constant". But this is not in accordance with common practice which is to define K as a constant of the acid-base system and to give it a value measured under specified standard conditions, which include the temperature and the solvent.

3. The activity coefficients might be combined with $(B)/(A)$ so that we should write

$$\log K = \log (H^+) + \log \frac{(B)}{(A)} \cdot \frac{f_{H^+} \cdot f_B}{f_A} \tag{9}$$

There is no objection in principle to this, but it is not very useful. Where an indicator or buffer is used, it is the ratio $(B)/(A)$ that is measured or fixed; where the behaviour in this solution of another electrolyte is considered, the activity coefficients of its ionic forms and of those with which it interacts are required for the description, not those of B and A. The third term of equation (9) does not therefore supply all the information that is required for such description and equation (9) offers no advantage as compared with equation (4).

4. The activity coefficients might be combined with $\log(H^+)$ to give

$$\log K = \log (H^+) \cdot \frac{f_{H^+} \cdot f_B}{f_A} + \log \frac{(B)}{(A)}$$

or

$$pH = pK + \log (B)/(A) \tag{10}$$

where

$$pH = -\log (H^+) \cdot \frac{f_{H^+} \cdot f_B}{f_A}$$

This also is theoretically unobjectionable. It might be positively justified if it could be shown that this function is given directly by the practical means for measuring pH. We shall proceed to examine these.

There are two general means of measuring pH: by the observation of chemical equilibria and by measuring the EMF of an electrolytic cell.

1. *Chemical equilibria.* There are three methods of measuring or fixing pH which all yield the same sort of function of concentrations and activity coefficients.

a. A buffer is used to fix the pH by fixing the value of the ratio $(B)/(A)$ in equation (4). Given the value of K of the buffer, the pH obtained directly is that of equation (10).

b. An indicator is added to a solution whose pH is fixed but unknown and the ratio $(B)/(A)$ of the acid and basic forms of the indicator is measured. Given the value of K of the indicator, the pH obtained is that of equation (10).

c. The rate of a hydrogen ion-catalysed reaction is measured. This is determined by the concentration of the complex A formed between the substrate B and the hydrogen ion

$$H^+ + B \rightleftharpoons A$$

The ratio of rates in two solutions thus measures the ratio of concentrations of A in them: the pH difference measured is that of equation (10).

2. *Potentiometric methods.* There are two sorts of cell used:

a. Cells with liquid junction. The type of cell most often used consists of a hydrogen or glass electrode dipping into the solution, with a calomel reference electrode

connected with the solution by a salt bridge; the "liquid junction" occurs between the salt bridge and the experimental solution. This type of cell has the great disadvantage that, in whatever way the liquid junction is made, the processes at it are irreversible and are not therefore subject to thermodynamic theory. Guggenheim (6) concluded that their EMF can be described only in terms of the mobilities and concentrations of all the ions present in all parts of the cell and our present knowledge does not allow this to be done more than very roughly. Obviously such a cell can give no theoretically satisfactory measurement of pH.

b. Cells without liquid junction. Cells of this type consist of two reversible electrodes dipping into a single solution. Their EMF can be precisely described in thermodynamic terms. Such a cell would consist of a hydrogen or glass electrode with, perhaps, a silver-silver chloride electrode reversible to Cl^-.

$H_2 \cdot Pt$/Solution containing buffer and neutral salts including H^+ and

$$Cl^-/AgCl \cdot Ag$$

The EMF of this cell at standard pressure of hydrogen gas is given by

$$E = E_0 + RT/F.\, ln(H^+)(Cl^-) \cdot f_{H^+} \cdot f_{Cl^-}$$

where E_0 is the EMF of the cell when it contains HCl of activity defined as unity. Since the concentration of Cl^- can be estimated chemically, one might define pH as measured by the EMF of such a cell containing unit concentration of Cl^-.

$$\text{pH} = -\log(H^+) \cdot f_{H^+} \cdot f_{Cl^-} \tag{11}$$

The reference electrode might be reversible to some other cation such as Na^+; then its EMF is given by

$$E = E_0 + \frac{RT}{F} \ln \frac{(H^+) \cdot f_{H^+}}{(Na^+) \cdot f_{Na^+}}$$

and one might define

$$\text{pH} = -\log(H^+) \cdot f_{H^+}/f_{Na^+} \tag{12}$$

From the foregoing discussion it appears that none of the definitions of pH is uniformly satisfactory. Those that are obtained directly from a method measurement (equations 10, 11 and 12) have the disadvantage that the quantity measured contains not only a quantity characteristic of the solution, but also quantities which are characteristic of the measuring system. Thus different measuring systems applied to the same set of solutions will give different results. For example, two indicators will in general record different differences of pH between the same two solutions and this will be especially marked if the indicators are of different ionic charge type; two different buffer systems, especially if they are of different ionic charge type, will suffer different changes of pH with change of ionic strength; two cells without liquid junction which have different

reference electrodes will respond differently to the same change of ionic strength. Nor can it be argued that the activity coefficients which refer to the ions of the measuring system will be valid for another system in whose dissociation we may be interested; this will be true (and even then only approximately) only if this system is of the same ionic charge type as the measuring system and only in dilute solutions.

Thus if it is desired to give an exact description of the state of dissociation of a given solute in a given solvent mixture, in relation to a measured property of the solvent, it is necessary to measure or compute both the activity coefficients relevant to the measuring system and those relevant to the solute. No advantage is to be gained by an attempt to reduce equation (4) to three terms and it is better to let the activity coefficient term stand explicitly alone. This means a return to Sørensen's original definition of

$$\mathrm{pH_S} = -\log(H^+)$$

which is the only function characteristic of the solvent mixture and independent of the solute or measuring system. This is distinguished by the subscript S (denoting "Sørensen") from any empirical meaning of pH. It would prevent confusion if $\mathrm{pH_S}$ were used exclusively in would-be exact physicochemical expressions and reserved for this purpose.

At the same time "pH" as commonly used is very useful within certain limits. It does provide an easily made measure of the physico-chemical state of a system which enables that state to be approximately reproduced from one experiment to another. In solutions fairly dilute with respect to electrolytes and in which specific interaction of ions is small, the activity coefficients of different salts and uncharged species do not differ very much from each other or from unity and the estimates of pH of such solutions obtained by different methods—even from cells with liquid junction—will not be widely different; this is true even if it is difficult to interpret the measurement in terms of exact physico-chemical quantities. In quoting such a pH value it is important to state the method by which it was obtained; for example, the buffer mixture or indicator, the assumed value of pK, the concentration of neutral salt, the electrode system and the standard of pH used to calibrate it. Indeed, it is often more precise and no more trouble to state the composition of the solution alone; for example, to state that a protein was dialysed against "$NaCl$ 0.2M, Na_2HPO_4 0.1M, KH_2PO_4 0.1M" rather than "0.2M phosphate pH 6.8 containing 0.2M $NaCl$".

To summarise, it is highly desirable that pH should be confined to two meanings:

1. The ordinary empirical quantity used as an approximate specification and derived directly from a method of measurement; in quoting its value, all the relevant details of the method should be given. It should be written "pH".

2. In any exact physico-chemical expression where its use is desired for brevity, it should mean only $-\log(H^+)$ and should be written "$\mathrm{pH_S}$". Activity coefficients should appear explicitly as a separate term, as in equation (4).

II. *Bulk Phases in Membrane Equilibrium.* Bulk phases separated by a membrane which allows the passage of some but not all of the ions deserve some special discussion; differences in their properties are of interest in accounting for the differences of composition across natural membranes, for example, between the interior of a cell and its surrounding fluid. Since each of two phases considered is a bulk phase, there might seem on the one hand to be no need to add to the foregoing discussion; on the other hand, since they are in equilibrium, it might be tempting to give a thermodynamic rather than a stoichiometric meaning to pH and to say that it ought to be taken to be the same in both the phases.

However, it is not legitimate to use any such concept as "activity of hydrogen ion", which can be measured or defined only relative to some arbitrary definition. What can be measured is either physically indefinite or involves other ions as well, depending on the method of measurement. Thus, as in the case of single bulk phases, the values obtained depend on the method of measurement and the activity coefficients which their expression involves are relevant only to the particular method used.

1. *Equilibrium methods.* Owing to the Donnan equilibrium effect, the distributions of ions of different charge type will differ. Thus the values of $(B)/(A)$ for a buffer or indicator will be different in the two phases; the values will depend also on the ionic radii and on factors leading to specific interaction between ions. The ratio of the values of $(B)/(A)$ found in the two phases with a given indicator are the same whether the indicator species are themselves in equilibrium across the membrane, provided that they can pass through it, or whether small amounts of indicator are added to separated samples of the phases. A solution obtained by dialysing a charged colloid against a buffer is not identical with that obtained by dissolving the colloid in the same buffer, even if the colloid is first adjusted so that it will not appreciably exchange ions with the buffer system.

2. *Electrode systems.* The passage of current between two electrodes, one being in each phase can lead only to:

a. The transfer of combinations of ions satisfying the condition of electrical neutrality (5), if the electrodes are reversible to ions that can pass the membrane.

b. To polarisation of the membrane if the electrodes are reversible to ions that cannot pass the membrane.

c. More complex changes if the electrodes are not reversible.

Of these, *b* will give no useful information while *c* will be difficult to interpret. *a* is represented, for example, by a pair of hydrogen or silver chloride electrodes, one in each phase; the potential difference between such a pair is zero. Similarly, the potential differences between a reversible hydrogen and a reversible chloride electrode, both in the same phase, are the same in each phase. These identities correspond with the identity of the activity of any combination of ions, that satisfies the condition of electrical neutrality, in the two phases. On the other hand, the potential difference between two calomel-KCl electrodes, one in each phase, is not zero; nor are the potentials of hydrogen against calomel-KCl measured in the two phases the same; the reason for these differences must

lie in the irreversible nature of the processes that occur at the liquid junctions of the calomel electrodes.

There seems to be no reason, therefore, to vary the conclusions reached for the case of single bulk phases. In using pH in its empirical sense, it becomes even more important to specify the conditions of its measurement. For exact expressions the use of pH_S is desirable, which leaves the effects of activity coefficients and of Donnan distribution to be expressed explicitly.

Membrane potentials. Appeal is often made to an electrical potential difference across the membrane to explain the unequal distribution of given ions. The difference of potential between two calomel-KCl electrodes, one in each phase, is often taken as measuring this potential. Since use is made of this potential in connection with observations on the osmotic pressures of solutions of charged electrolytes and their "pH", the matter requires some attention.

The derivation of the Donnan distribution does not require the assumption of any membrane potential; only two conditions need be imposed to obtain it:

1, that both phases shall be electrically neutral;

2, that the activity of any combination of ions, whose transfer across the membrane involves no net transfer of charge, is the same in both phases (e.g., a pair of oppositely charged univalent ions).

All the properties of the system, including the osmotic pressure, can be expressed in terms of the stoichiometric concentrations of ions in the two phases and the activity coefficients of salts and there is no need to assume any value of potential at the membrane to account for any of these measurable quantities.

There is no doubt that there does reside in the membrane a discontinuity to which the name "potential" can be given. But any definition of its magnitude must involve the work of transfer of a single sort of ion from one phase to the other; this involves an assumption about the activity coefficients of that ion in the two phases which, as has been seen, must be arbitrary. The case is exactly that of the potential of a single electrode against solution. The stoichiometric ratios of concentrations of different ions in the two phases will in general be different so that there can be no unique standard even of arbitrariness. The membrane potential could be used only to give an account of the stoichiometric ratios, but it is preferable to deal with these directly and explicitly. The membrane potential is, therefore, a useless abstraction.

The quantity measured as the "membrane potential" is usually the potential difference between two calomel-KCl electrodes, one in each phase. While this is a reproducible quantity, under given conditions of experiment, its measurement involves the use of two liquid junctions of KCl solution against solutions of different composition. The assumption that the whole—or any definite part—of the observed potential difference arises at the membrane is completely arbitrary and any attempt at testing its truth must involve an arbitrary assumption about the activity coefficients of some one sort of ion.

The use of the concept of "membrane potential" is therefore to be avoided. If any assumption has, for convenience, to be made, it is much better based on a ratio of concentrations than on measurements with calomel electrodes.

III. *Interphases*. The state of dissociation of an acid group in an interphase is often different from the state of dissociation that it has, or from its chemical nature would be thought to have, in a bulk phase in equilibrium with the interphase. Such a difference is of interest to the biologist in connection with the properties of proteins and other colloids and of cells. It is desirable that such differences should be expressed in satisfactory terms.

Interphases differ in some important respects from bulk phases. They are the regions of discontinuity between bulk phases. They are very small in bulk and are anisotropic. Care is needed that terms applied to them shall be physically definite and can be related to measurements that can actually be made on them. Thus, to consider, as Craxford, Gatty and Teorell (7) do, the readings that might be obtained with electrodes placed in the interphase is not helpful, because it would not be possible to know in practice in what part of the interphase the electrodes had been placed nor what disturbance they might have introduced; nor can analytical measurements be made on an interphase. The only properties of an interphase that can be measured are those relative to the bulk phases; for example, interfacial tension (that is, the change of energy of the system on increasing the area of interphase), electrophoretic mobility or streaming, or the changes in the composition of bulk phases that result from the formation of interphase. The state of dissociation of groups in an interphase can be deduced from measurements such as the change of interfacial tensions over a solution of a fatty acid or amine with pH of the bulk phase (Peters) (8), the change of electrokinetic properties with change of composition of the bulk phase (Danielli) (10) or the changes in the colours of indicators in the presence and absence of interphase.

Without for the moment specifying the meaning of its terms, acid-base equilibrium in an interphase may be stated to be expressed by an equation analogous to (4). In comparing such an equation with that applying to the bulk phase, special care will be needed about the meaning of "concentration". Only the Gibbs "surface excess" Γ of any molecular species can be measured; this may be positive or negative and its difference from zero certainly implies a difference of something resembling "concentration" between interphase and bulk phase. But the interphase is anisotropic and even an average concentration cannot be obtained from Γ unless the thickness of the interphase is known. The thickness is usually unknowable and therefore subject to arbitrary definition; only in certain rather simple cases can any quantity be obtained which is in any way analogous to thickness. These are the cases where the A and B forms of the substance or group in whose dissociation we are interested are materially confined to the interphase; that is, in which the surface excesses of A and B are very large and positive, compared with their concentrations in the bulk phase. Such cases are those, for example, of monolayers of long-chain fatty acids or the surface of particles of protein. The degree of dissociation in the interphase can then be sharply distinguished from that in the bulk phase, since the bulk phase contributes negligibly to the quantities of A and B in the interphase; the other particles in whose distribution we are interested are then ionic and their dis-

tribution is largely determined by electrical forces arising in the interphase. We shall confine attention to interphases of this sort.

Now consider equation (4). Since A and B are materially confined to the interphase and only their ratio enters the equation, these can be defined in any convenient way, such as quantities per unit area of interphase. The activity coefficient term is dimensionless. It follows that K and (H^+) are of the same dimensions. Therefore, if K is to be of the same dimensions as in the bulk phase, then (H^+) must have the dimensions of a volume concentration; or if (H^+) is to have other dimensions, such as a surface excess, then K must be correspondingly defined. But, in the latter case, the value of K would not be comparable with that in the bulk phase and it would be necessary to refer all thermodynamic quantities to a standard state defined for the interphase, which would be experimentally inconvenient. We shall therefore keep K having the dimensions of a volume concentration.

In terms of equation (4), the observed differences of $(B)/(A)$ in interphase and bulk phase could be described in three ways, which we shall discuss in turn:

1. *Difference of K.* One would expect the tendency of an acid molecule to dissociate to be different in an interphase and a bulk phase; not only may there be an electrical potential gradient in the interphase, but other physico-chemical conditions also may be considerably different, such as dielectric constant and the basic strength of the solvent. Hartley and Roe (9) suggested a difference of K as a means for expressing the difference of degrees of dissociation in interphase and bulk phase in equilibrium. This is certainly permissible and might at first sight appear particularly reasonable. But it must be remembered that K may be given any numerical value whatever according to the definition of a standard state. A rational definition of K in the interphase would be with reference to some accessible standard state in the interphase, but this has the disadvantage which has been mentioned that all measurements on an interphase are in fact made relative to a bulk phase. Hartley and Roe choose as a standard state in the interphase that in which there is no gradient of potential; while this allows the Debye-Hückel treatment to be applied to the variation of K with composition of the bulk phase, the standard state cannot be even closely approached in many cases. In view of the fact that, in any case, the numerical values of K are subject to the definition of the standard states and therefore do not measure the difference of the tendencies of an acid to dissociate, it seems simplest to retain the bulk phase value of K for use in the interphase.

2. *Difference of (H^+).* There will certainly be a surface excess of hydrogen ion in the interphase, positive or negative, according to the electrical and other physical properties of the interphase. Insofar as it depends on the electrical properties, the Debye-Hückel theory allows this surface excess and its distribution in depth of the interphase to be calculated in simple cases. The influence of other physical factors cannot be allowed for, so that a calculated value of the surface excess and its distribution has no absolute significance; it allows the variation of the properties of the interphase with change of composition of the bulk phase to be predicted, insofar as these change the electrical properties of the

interphase. Now, (H^+) must have the dimensions of a volume concentration, whereas the distribution of hydrogen ions in the depth of the interphase is non-uniform. Hartley and Roe (9) calculate a differential value for (H^+) as a function of the charge of the interphase and the ionic composition of the bulk phase; this corresponds physically with the quantity of hydrogen ion per unit volume of an infinitely thin lamina infinitely near to a surface in which the charge is located. This definition of (H^+) is dimensionally correct and physically exact; it is a perfectly legitimate means of expressing the variation in interphase properties with those of the bulk phase.

Danielli (10) also calculates (H^+) by the much less legitimate device of supposing that there is a Donnan membrane parallel with the charged surface and distant $1/\kappa$ (the characteristic distance of the Debye-Hückel theory) from it; he assumes uniform concentrations on both sides of this membrane and Donnan equilibrium across it. He is fortunate that, for surfaces of small curvature, this treatment yields the same value of (H^+) as does Hartley and Roe's.

The objection to the use of the Debye-Hückel theory to calculate a value of (H^+) in the interphase is that it is not in accordance with the standard application of it to small ions. Hartley and Roe suggest that it should be applied to the calculation of the value of (H^+) at the surface of even small ions and that the value of K for such ions should be that in which there is no surface excess of hydrogen ions. But such a state is in practice unattainable and there seems to be no good reason for abandoning the traditional use of the Debye-Hückel theory for calculating activity coefficients relative to an approachable standard state.

3. *Activity coefficients.* The use of the activity coefficient term has the advantages that it is the traditional method of treating the variation of the properties of small ions and that it introduces no new difficulty over the definitions of K and (H^+), which have the same meanings and values as in the bulk phase. The activity coefficient expresses in a formal fashion differences in the mutual energies of ions, as measured or calculated, from their values in a defined and accessible standard state. The illegitimacy of the expression of the activity coefficient of a single ion is well known. The setting up of physical models which have a spurious air of reality, without giving any experimental or theoretical advantage, is avoided. Cannan (11) has made use of the calculation of activity coefficients in describing the titration curves of complex electrolytes as a function of the composition of the solution.

All these three involve the same application of the Debye-Hückel theory and differ only in the way in which the results are expressed. The choice is therefore not between principles but between modes of expression. The last is to be preferred. The term "pH" has been intentionally avoided in the discussion of interphases. It is now seen that "surface pH" can refer only to a surface excess or to a function of a rather unreal "concentration". By the proper use of activity coefficients its use can be avoided. Loose talk about the "pH of interphases" is to be condemned.

Finally, it should be remembered that only the electrical properties of interphases can at present be treated theoretically in a quantitative manner, and

then only in simple cases. The changes in degrees of dissociation of groups in the interphase with changes of composition of the bulk phase can be so far predicted, and are reflected in the changes of measurable properties of the interphase relative to the bulk phase. Agreement between the changes of several measurable quantities does not necessarily strengthen confidence that no factors other than the electrical are of importance. For example, the agreement between thermodynamic and electro-kinetic changes found by Danielli (10) serves only to show the consistency of the Debye-Hückel treatment as applied in these two fields, not the correctness of a predicted "surface pH".

CONCLUSIONS

It is concluded and recommended that:

1. Wherever the term "pH" is used in exact thermodynamic equations it should bear the single meaning "$-\log(H^+)$" and should always be written "pH_S".

2. Where "pH" is used as an empirical means of specifying the acid-base properties of solutions, derived from some particular means of measurement, it should be written without subscript. In giving its value, the method of measurement, its standardisation, and details of the composition of the solution (such as the kinds and concentrations of buffer and neutral salts) should also always be specified.

3. The use of the concept of "membrane potential" should be avoided.

4. The term "pH" should never be used with reference to interphases.

REFERENCES

(1) Brønsted, J. N. Rec. Trav. Chim. 42: 718, 1923.
(2) Lowry, T. M. Chem. and Ind. 42: 43, 1923.
(3) Sørensen, S. P. L. Compt. rend. Lab. Carlsberg 8: 1, 1909.
(4) Lewis, G. N. and M. Randall. Thermodynamics. 1909.
(5) Guggenheim, E. A. J. Phys. Chem. 33: 842, 1929.
(6) Guggenheim, E. A. J. Phys. Chem. 34: 1758, 1930.
(7) Craxford, S. R., O. Gatty and T. Teorell. Phil. Mag. 25: 1061, 1938.
(8) Peters, R. A. Proc Roy Soc. A 133: 140, 1931.
(9) Hartley, G. S. and J. W. Roe. Trans. Faraday Soc. 36: 105, 1940.
(10) Danielli, J. F. Biochem. J. 35: 470, 1941.
(11) Cannan, R. K. Chem. Rev. 30: 395, 1942.

A REVIEW OF PHYSIOLOGICAL AND PSYCHOLOGICAL STUDIES OF SEXUAL BEHAVIOR IN MAMMALS

FRANK A. BEACH

Department of Psychology, Yale University, New Haven

Purpose and Methods. As used in this article the term "sexual behavior" refers exclusively to the overt acts comprising heterosexual copulation, to those contiguous reactions commonly designated as "courtship" or "precoital play," and to a variety of non-copulatory sexual responses such as those involved in auto-erotic and homosexual activities.

The references and discussion which follow have been assembled and organized in an attempt to serve three purposes: (1) to bring together various neurological and endocrinological data concerning the strictly physiological aspects of sexual behavior; (2) to integrate with these relatively clear-cut results of laboratory investigation certain less definitive but equally important psychological findings reported by students of animal behavior, by practical animal breeders and by clinicians; and (3) to attempt to derive from this psycho-physiological potpourri a series of inter-species comparisons which may implement an evolutionary interpretation of human sexuality.

Merely to list and summarize the physiological experiments would have been a comparatively simple task, whereas the interpolation of non-physiological evidence has proven very difficult, and the result leaves much to be desired as far as logical organization is concerned. Nevertheless in preference to an annotated bibliography I have chosen to offer a complex and unavoidably incomplete survey of an important field of knowledge in the hope that presentation of the physiological and psychological data side by side will emphasize the essential unity of the phenomena dealt with by different disciplines, and with the expectation that the frequent lacunae revealed by this method may call attention to the urgent need for further research by investigators in many areas of science.

Importance of the Holistic Approach. Physiological experiments designed to identify the nervous pathways involved in a particular genital reflex, or to measure the importance of secretions from a single gland to the occurrence of copulatory reactions have contributed a great deal to our understanding of sexual behavior. It should be obvious, however, that the full significance of such findings becomes apparent only when they are viewed against the broader background of the total sexual pattern as it appears in the normal animal.[1]

Mating reactions are best interpreted as responses to internal and external stimuli,—and these responses are mediated by complex neuro-muscular mecha-

[1] It has been shown, for example, that decorticate female rats are capable of fertile copulation with the male (44), and this observation taken alone might be thought to signify that the loss of neo-cortex is without effect upon the female's sexual reactions; but subsequent investigation revealed that after destruction of the neo-pallium the mating pattern is modified so profoundly that males discriminate against operated individuals and will mate with them only when normal females are not available (45).

nisms whose functional capacity depends upon genetically-determined structure, upon concurrent activity in other nervous circuits, upon the composition and concentration of hormones present in the organism, and in varying degree upon the previous experience of the individual. None of these contributing factors operates in isolation, and the functions of one cannot be considered save in terms of its interrelations with all other variables involved.

FUNCTIONS OF THE NERVOUS SYSTEM IN SEXUAL BEHAVIOR. For purposes of analysis the rôle of the nervous system in sexual behavior may be considered as divisible into receptor and effector functions. In connection with both categories one must include the task of integrating separate units into a complete pattern. Thus, both the peripheral reception of separate stimuli and the central correlation of multi-sensory impulses belong under the same rubric; and motor functions are assumed to include not only the mediation of separate genital and postural reflexes, but the organization of these individual reactions into a unified pattern of response.

A. *Sensory Functions.* 1. *Evolutionary importance.* Identification of the sensory basis for sexual arousal is a matter of prime importance not only to the physiologist and the psychologist but to the student of evolution as well; and the concept of "sexual isolation" is one which systematists and geneticists find very useful in explaining the continued separation of closely related and inter-fertile sympatric species.

Sometimes referred to as "psychological isolation," this condition is defined as "absence of mating due to lack of mutual attraction" (98). "Sexual isolating mechanisms" are believed to include differences in behavior, in smells and in adornments which make males and females of different species less apt to find each other or less attractive and thus less willing to court and mate (99, 209).

Thus the problem of defining the stimuli responsible for sexual behavior is at the same time one of determining the basis for "species recognition" which is a governing force in evolutionary change.

2. *Inadequacy of simplified concepts.* Before one can appreciate the significance of experiments dealing with the functions of sensation in sexual behavior it is necessary to recognize the inadequacies of one particular line of thought which for many years tended to dominate research in this field and to obscure the true nature of the phenomena concerned. Over-simplified concepts of the neuro-physiological processes underlying sexual arousal, and naive interpretations of mating behavior as a sequence of concatenated reflexes early led to a search for *the stimulus* which set off the hypothesized chain reaction.

Presently available evidence permits the categorical statement that there is no single form of stimulation which can be regarded as the *sine qua non* for sexual arousal and copulatory reactions in male or female mammals (40). On the contrary, activation of several sensory systems normally contributes to the occurrence of the behavior in question and any single afferent pathway is dispensable provided the rest remain. From the psychological point of view it appears that the stimulus which initiates mating is not a single characteristic, but a pattern with certain characteristics of organization (187).

It is possible to go further and point out that even in the case of such restricted and incomplete segments of the total pattern as erection in the male or assumption of the copulatory crouch by the female the problem of sensory control is not a simple one; for almost every such reaction appears subject to alternative and supplementary forms of effective stimulation.

3. *Importance of individual differences.* It should be stated explicitly that investigation of the sensory control of sexual arousal involves problems of a higher order than those presented by the identification of stimuli adequate to elicit the separate muscular and glandular reactions involved in copulation. Definition of the stimuli evoking ejaculation is an important step toward the complete understanding of sexual activity; but exposition of the forms of stimulation responsible for pre-coital courtship and subsequent copulatory reactions is a much more complex affair which is likely to succeed only if approached with full recognition of the sources of variability capable of affecting the experimental results.

To illustrate,—male dogs and cats display erection in response to manipulation or electrical stimulation of the penis. This reaction can be evoked repeatedly and there seems to be little individual or temporal variability in the stimulus-response relationships involved. In contrast there are great differences between individual males or between the reactions of the same male at different times in respect to the ease and rapidity with which courtship and copulatory reactions can be induced.

Even when the conditions of external stimulation are held as nearly optimal and constant as possible the facility with which mating reactions are called forth differs markedly among male rats (4, 41), rabbits (196), guinea pigs (218), cats (reviewer's observation), sheep (211), pigs (144), rhesus monkeys (175), chimpanzees (55, 105, 326) and, I suspect, every other mammalian species. Individual differences in sexual responsiveness appear in male rats which have been reared in isolation (41), and are so extreme that a few obviously healthy and fertile males never attempt to mate with the receptive female (4). Variations in sexual excitability occur in boars despite a normal hormone balance (144), and wide individual differences are seen in male rats castrated at birth (50).

Differences in the ease with which initial arousal occurs are an important factor in any investigation of sensory function in sexual behavior, for they determine in part the quantity and quality of external stimulation necessary to evoke copulation. Although the majority of sexually-inexperienced male rats attempt to mate only when stimulated by a female in full estrus, a few highly excitable individuals which have been reared in isolation execute copulatory attempts when presented with non-receptive females or young males (41). Most male baboons actively court only those females in which the sex skin is swollen, but some males repeatedly mount non-estrous females (335); and in like fashion some male chimpanzees try to mate with any female at any stage of her sex cycle, while other males will copulate only at the height of the female's genital swelling and may even display a strong preference for individual females (105, 326). In general

it appears that, other factors being equal, "the specificity of the stimulus adequate to elicit mating responses varies inversely with the sexual excitability of the individual" (40).

Although presently available evidence would appear to suggest that the magnitude of individual differences in sexual responsiveness is much greater in male than in female mammals, there are several indications that females differ one from another in the ease with which copulatory reactions may be evoked. Some female rats in heat display lordosis each time they are mounted by the male, whereas other estrous females do not exhibit this reaction every time they are stimulated but do so in response to a constant and predictable proportion of the male's copulatory attempts (45). Similarly, ewes in full estrus show individual differences in willingness to stand for the ram and in readiness to take the initiative in courtship (211).

Individual differences in sexual excitability among male animals not only affect the initiation of copulatory behavior, but may be seen to influence the amount of stimulation necessary for ejaculation. Some male rats reach the peak of excitement essential to ejaculation after only two or three intromissions, whereas other males of the same strain do not attain this level of arousal until intromission has occurred fifteen or twenty times (35, 314). It is especially significant to note that in repeated sex tests the frequency of copulation in a given period and the number of intromissions preceding ejaculation are relatively constant for individual male rats as shown by the high reliability coefficients reported by different investigators (4, 31). Observations of copulatory behavior in the rhesus monkey have revealed that the number of mounts preceding intromission may vary from three to one hundred depending upon the "virility" of the individual male (75).

It seems probable that variations in sexual excitability are responsible in large measure for marked differences in the general character of the sexual approach shown by different males, particularly those of the higher primate species. Some male chimpanzees react to the female very abruptly, indulging in little or no foreplay and copulating vigorously or even violently, whereas other males practice a great deal of pre-coital stimulation, treat the female gently and copulate only after considerable preparation (55).

4. *Functions of non-contact stimulation.* a. *In the female.* The limited experimental evidence might lead one to believe that mating behavior of female mammals normally takes place without the intervention of visual, auditory or olfactory sensations; for removal of the eyes and the olfactory bulbs and destruction of the cochlea fails to abolish copulatory responses in the female cat (23) or rabbit (65). Actually, of course, these findings mean merely that when such drastically desensitized females are mounted by the male they respond to the resultant contact stimuli with appropriate postural adjustments facilitating intromission.

The complete sexual pattern of estrous females of most mammalian species includes active orientation to the distant male, and if the male is slow to respond

the female may initiate courtship. For example, the onset of estrus in the heifer is signallized by her tendency to seek out and stand beside the bull (1a), and female rats in heat display an increased drive to cross an electrified grid interposed between them and the male (313). If the male fox is not sufficiently aggresive the estrous vixen begins courtship play (108). Although the non-receptive female rhesus usually avoids contact with the male, when in heat she repeatedly approaches him, displaying a characteristic stance and stereotyped gestures (75). Female chimpanzees at the height of genital swelling may exhibit the sexual presentation and by various gestures and vocalizations invite the male to copulate (326).[2]

Reactions such as these obviously depend upon distance reception of one type or another, and although supporting experimental evidence is lacking it seems highly probable that the behavior described would be lacking in females which were blind, deaf and anosmic.

b. *In the male.* Although female mammals generally show an increased tendency to associate with males during the breeding period and may even instigate courtship when the masculine partner is slow to respond, it is commonly the male who plays the more aggressive and dominant rôle in the mating relationship. Even when the female begins the foreplay, initiative eventually must shift to the male if copulation and fertilization are to be accomplished, and it may be for this reason that distance receptors play a more fundamental part in the sexual functions of male than of female mammals.

The male animal's ability to detect the female's readiness to mate varies from species to species, from individual to individual, and even from time to time in the same individual. In the case of some sub-mammalian forms, such as certain amphibia, there is little indication that males can discriminate between receptive and non-receptive females, and sex recognition seems to be largely a matter of trial-and-error (228). Even male mammals which ordinarily are capable of discriminating between the estrous female and other animals may fail to respond selectively when highly aroused. Thus male rats (289) and guinea pigs (270) excited by recent contact with the receptive female often attempt to mate with other males or with diestrous females; and buck rabbits abruptly separated from the estrous doe mount and ejaculate in response to a dummy female (103) or even the hand of the experimenter (196). Under similar conditions male dogs mount males or the non-receptive bitch (129), and male rhesus monkeys may mount other males, sometimes with anal penetration (76).

Despite occasional absence of discrimination there is ample proof that prior to physical contact males of many mammalian species respond differentially to other males, non-receptive females and females in heat. Male rats in revolving cages become more active if cages containing estrous females are placed nearby

[2] In the majority of human cultures open sexual solicitation by unmarried females is socially forbidden; but in some primitive tribes which constitute exceptions to this rule the female's invitation takes the form of direct exposure of the genitals (122), a reaction strongly reminiscent of the sexual presentation shown by females of other anthropoid species and of many lower primates as well.

(45 cm. or less) (269).[8] The male chimpanzee plainly recognizes when the female in an adjacent cage is sexually receptive, as his prompt erection and invitational gestures amply convey. Observation suggests that in this case the effective sensory cues are the female's gross movements, gestures, facial expression, and the condition of her external genitalia (326).

Results of experimental studies suggest that in male lagomorphs and rodents several distance receptors probably contribute to sexual arousal and to the orientation of mating activity. Removal of the eyes or the olfactory bulbs does not eradicate copulatory behavior in sexually experienced male rabbits (65, 262, 290) or rats (41, 286, 287); but either operation performed prior to the acquisition of mating experience reduces the probability of copulation in male rats (41). The finding that blindness or anosmia tends to increase the proportion of rats which fail to copulate has been interpreted as evidence of lowered sexual responsiveness consequent to the reduced sensory intake (40).

The importance of olfactory stimulation to sexual activity of male animals has been greatly exaggerated by some workers. It has, for example, been claimed that male rats isolated for six months from the odor of females suffer testicular atrophy and complete loss of sex drive, and that such animals are restored to physiological normality when females are placed in nearby cages (278). Repetition of the original experiments with adequate controls and careful study both of behavior and histological condition of the gonads and accessories has shown that none of the reported changes take place in males isolated for eleven months (239). Similarly the arrested genital development and reproductive failure which is reported to follow removal of the olfactory bulbs in young male guinea pigs (199) probably is due, not to the loss of olfactory sensation, but to the severe brain infection which often follows this operation.

Laboratory findings concerning lower mammalian forms appear to minimize the functions of olfactory stimulation in the male's sexual arousal, whereas observational evidence has led commercial animal breeders and field naturalists to conclude that the odor of the female in heat is a potent force attracting and exciting the male. The validity of this conclusion has not yet been experimentally established but it seems probable that sensations of smell do contribute materially to the processes under consideration, although they are not, in all likelihood, the sole source of sexual stimulation in any species.

When they are in estrus the females of several mammalian species including the house cat (214), African lion (88), horse and cow (248) utter characteristic vocalizations, and it has been suggested that these "sex calls" stimulate males of the same species. Here again, although the supposition could be tested experimentally it has not been, and the only laboratory data available are those

[8] Although these findings originally were taken to indicate that the males sensed the presence of females in heat, the data are susceptible to an alternative interpretation. Female rats exhibit a striking increase in activity at the time of estrus, and it is possible that increase in the male's activity was simply a response to auditory and vibratory stimuli provided by the revolutions of the female's cage.

indicating that partial destruction of the cochlea does not eliminate mating in the male rat (286, 287) or rabbit (65).

Mating behavior survives combined elimination of visual and olfactory end-organs in the experienced male rabbit (65) and rat (286, 287); but copulatory reactions rarely appear in male rats which are rendered blind and anosmic before contact with the receptive female (41). The fact that the combination of blindness and anosmia depresses sexual responsiveness to a greater degree than does either type of sensory loss alone is taken as additional proof of the theory that both visual and olfactory stimuli normally contribute to sexual arousal (40).

5. *Functions of contact stimulation.* a. *Extra-genital stimulation in the female.* Various types of extra-genital contact elicit the copulatory posture and appear to increase sexual excitement in estrous female mammals. Some female rats in heat exhibit hopping, ear-vibration and lordosis when the back and sides are stroked or palpated lightly with the thumb and forefinger (16, 314). Stroking the rump of the estrous female guinea pig evokes opisthotonus and elevation of the perineum (329), and similar reactions may be elicited in the sexually receptive golden hamster by stroking or pressing the small of the back (reviewer's observation). If one presses heavily upon the back and sides of the sow in heat and simultaneously moves the tail to one side, the female responds by lifting the tail and hind quarters, presenting the genitalia, and grunting as in copulation (3). Receptive female cats can be induced to assume the mating position if the experimenter grips the loose skin of the animal's neck between his fingers and strokes or presses downward upon her back (23).

In each of the foregoing situations the inciting stimuli mimic those normally provided by the reactions of the sexually active male. Thus, prior to mounting the female, the male cat grips the skin of the back of her neck with his teeth, and the resulting stimulation is partially responsible for the female's assumption of the copulatory crouch,—a position essential to complete mating. A similar function should probably be referred to the comparable mating grip which is shown by males of several other species, including the short-tailed shrew (232), marten (204), mink (109) and bat (317). In some species such as the bat the male does not show penile erection until after gripping the female.

In the normal pattern of pre-coital play males of many mammalian species apply a variety of forms of tactual stimulation to the receptive female. In most species the male bites or nibbles at many parts of the female's body but animals with prehensile forelimbs may employ the hand and fingers as well as the mouth. The male monkey's manual grooming of the female increases in frequency when she is in heat (75), and involves "stimulation of most of the major receptor zones" (175).

Male animals of various species often bite or suck the female's ear before or during coition, and although there is no physiological evidence to prove that this form of stimulation is sexually exciting to the female, her subsequent responses often suggest that such is the case. In this connection it is significant to note that manipulation of the external ear produces contractions of the smooth muscles of the vagina in the estrous female cat (63).

The stimulatory activity of the sexually aroused male sometimes develops into active aggression with the result that the receptive female may suffer severe wounds. This is commonly observed in such widely separated forms as the marten (204) and the rhesus monkey (75).

Many of the pre-coital and coital activities characteristic of other animals are also present in human behavior, and the high frequency of such common elements suggests the existence of a basic pattern underlying the sexual activities of all mammalian species. For example, although no reliable evidence has been published to indicate the types of sexual stimulation most frequently employed in European cultures, it is well established that in many primitive tribes such as the Trobriand Islanders (201) and the Ainu and Toda of Eurasia (122), biting, scratching and hair pulling are common precursors and accompaniments of intercourse.

Despite the impressive degree of inter-species coincidence in many sexual responses it is obvious that specialization of morphological structures or nervous arrangements may give rise to particular forms of behavior which are restricted to the species so affected. It is well known, for instance, that in the human female the breast is a particularly sensitive erotic zone, and manual and oral stimulation of this area is a prominent element in the pre-copulatory pattern not only in European cultures but in other groups such as the Hopi of North America and the Lepcha of Eurasia (122). The "love magic" of certain Melanesian tribes involves surreptitious touching or caressing of the nipples of the woman whose favors are sought,—a custom based upon implicit recognition of the sexually-exciting powers of such stimulation (201). There is little evidence for this particular type of sexual stimulation in lower mammals, although one worker, observing that female rhesus monkeys in heat may manipulate and suck their own nipples, has interpreted this as auto-erotic (76).

b. *Extra-genital stimulation in the male.* Male rats continue to copulate despite surgical interference with cutaneous sensitivity of the entire ventral body surface (286, 287), and transection of the sensory nerves serving the snout and lips does not abolish mating activity in this species (41); but except for these reports there are very few data to indicate the importance of extra-genital sensations of touch and pressure to the occurrence of sexual arousal and coital behavior in male animals.

c. *Genital sensations in the female.* It was at one time a rather widely accepted view that sensory impulses from one or another part of the reproductive system are responsible for sexual arousal in female mammals; and one finds in the literature the statement that estrous changes within the vagina are, "the real internal stimuli for sex behavior in the female rat" (314), or that the "female hormones are psychologically important where they increase the local irritability of the sex organs, and thereby bring increased drive for further stimulation" (159a). A less extreme view is represented by the suggestion that sensations arising from direct vaginal or cervical stimulation are essential to and responsible for the receptive female's copulatory responses.

Several lines of evidence combine to indicate that neither of these theories is

sufficient to account for the behavioral phenomena under consideration. Normal estrous reactions occur in estrogen-treated female rats with congenital absence of ovaries, tubes, uterus and external vagina (48) and this would appear to negate the hypothesis that sensations from accessory sex structures are indispensable, inciting stimuli.

The belief that stimulation of the vagina and cervix is essential is equally untenable. In the first place, females of various species occasionally exhibit characteristic heat responses prior to any physical contact with a second animal. Highly receptive female rats may display hopping, ear-vibration and lordosis even though the male does not approach them (reviewer's observations), and some estrous cats crouch and tread spontaneously in the absence of the male (24). Secondly, as was pointed out in the preceding section, the copulatory adjustments of the estrous female are often elicited by various types of bodily contact which do not involve genital stimulation (e.g., crouching and treading may be evoked in the female cat by gripping the loose skin of the neck).

Finally, various experimental studies have shown that elimination of sensations from the genital region does not abolish mating behavior in the female. Removal of the uterus and vagina fails to eliminate copulation in the rat (11), and sexual reactions survive deafferentation of the entire genital region in the female rat (9), rabbit (65), and cat (21). Characteristic mating responses appear in female mice (247) and rats (48) despite congenital abnormality which includes absence of an external vagina.

Surgical removal of the tubes and uterus has even been reported to increase sexual responsiveness in approximately one-third of the women so operated (312), but this change may well be due at least in part to eradication of the fear of pregnancy.

Females of several species including the rabbit (66), cat (127), short-tailed shrew (232), mink and ferret (109) do not ovulate spontaneously, but usually do so only after copulation with the male. It has been assumed that in such cases mechanical stimulation of the vagina and cervix gives rise to nervous impulses which cause the hypophysis to release an "ovulation-inducing substance" which acts upon the ovary. It is, however, informative to note that when estrous female rabbits mount other females in masculine fashion ovulation is sometimes induced in the mounting individual, although under such conditions vaginal stimulation must certainly be minimal (158). Furthermore, female mink and ferrets in heat may ovulate after a prolonged struggle with the male despite the fact that mating is not accomplished and intromission does not occur (109). It would appear that in species which do not ovulate spontaneously ovulation is normally facilitated by the occurrence of sexual arousal and that intense excitement may call forth the ovarian response in total absence of primary genital contact.

If it is accepted that impulses from the genitals are not essential to the basic coital performance of female mammals it must be equally apparent that under normal conditions genital sensations provide one very important source of increased sexual excitement. In many species the male manipulates the female's

genitalia before coition; and among lower mammals such as the rat, guinea pig and golden hamster the male's oral investigations of the vaginal orifice often evoke overt responses on the part of the receptive female. The estrous hamster, for example, elevates the perineum (247b) and moves the hind quarters convulsively from side to side in response to the male's lingual exploration of the ano-genital region (reviewer's observations). When the male porcupine noses the vulva of the female in heat she backs toward him vigorously, pushing actively against the source of stimulation (265a). The male eastern skunk induces the receptive female to assume the copulatory pose by scratching at her vulva with his hind foot (315a).

Species with prehensile forelimbs usually employ manual as well as oral techniques of genital stimulation. The pre-copulatory play of rhesus monkeys involves mutual genital manipulation which "generates obvious excitement" (175). The male chimpanzee's manual and labial manipulations of the female's clitoris call forth erection and motility of this member, and female chimpanzees in heat have been observed to practice mutual cunnilinguis (55).[4]

Masturbation of the woman is common in the pre-coital play of various human tribes such as the Anunta of Oceana (122), and in certain sections of the Trobriand Islands both manual and oral forms of stimulation are characteristically employed (201). The effectiveness of such activity rests upon the innervation of the regions affected, and it is known that in the human female the clitoris and inner surface of the labia minora are richly supplied with sensory receptors for touch, although the vaginal entrance is relatively poor in sense endings of any type (53).

Evidence for the importance of genital sensations in feminine sexual behavior is not entirely inferential. Although rats surgically deprived of uterus and vagina are capable of mating behavior, in consecutive copulations such individuals typically display a decrease in responsiveness to the male which is probably traceable to absence of genital stimulation (11). Normal estrous females of this species are sometimes partially resistant at the beginning of a sex test but frequently display a marked increase in willingness to accept the male after one or two intromissions have occurred; and female cats which fail to crouch when gripped by the male may do so after the penis comes in contact with the vagina (25). Some female rats in heat exhibit lordosis when a pipette is inserted in the vagina (45), and comparable mechanical stimulation evokes crouching and treading in the receptive female cat (25).

During ejaculation the male rat deposits a copulatory plug within the vagina of the female, and the resultant stimulation causes a momentary relaxation of the cervix permitting passage of the sperm mass into the uterine cornua (57). The foregoing sequence of events is strongly suggestive of orgasmic function occurring in the female in response to internal genital stimulation.

d. *Genital sensations in the male.* Despite occasional statements to the con-

[4] Observations of this type are pertinent to an understanding of the biological basis for various human sex practices and cast considerable doubt upon *a priori* definitions of "normal" and "abnormal" forms of genital stimulation (28, 244).

trary (226) sensory impulses from the penis do not appear to be indispensable to sexual arousal because male cats seize the female, display erection and copulate vigorously despite surgical anesthetization of the entire genital region (255); and male rats continue to show sexual excitement after surgical procedures have rendered intromission impossible (50).[5]

Nevertheless it is obvious that genital sensations constitute an important element in the stimulus pattern responsible for complete sexual arousal and copulatory reactions in males. Female chimpanzees may induce erection in the male by handling the flaccid penis (55); and the male dog displays erection and ejaculation in response to genital manipulation (23). In the absence of normal genital stimulation overt sexual behavior may be profoundly modified. For example, adult male rats with infantile penes exhibit a high proportion of incomplete matings, relatively few intromissions and extremely rare ejaculatory responses; and similar behavior is seen in males with fully developed copulatory organs from which a segment of the penile bone has been removed (50).

The importance of direct genital stimulation to complete sexual performance in humans of either sex undoubtedly is subject to considerable individual variation; and there is in addition some indication of the existence of racial, or at least cultural, differences in this regard. Melanesian natives are reported to achieve orgasm only after protracted copulatory contact involving much more genital friction than is necessary for climax in white Europeans (201); and results of a current interview study indicate that male American negroes from low socio-educational strata are much slower than whites in attaining orgasm. However, negroes of high social status (as measured by education, etc.) are closely comparable to whites in terms of ejaculation time (176).

e. *Genital sensations as a potential source of copulatory movements during sleep in lower forms and erotic dreams in humans.* Genital reflexes sometimes occur in the apparent absence of any external stimulation. For example, full erection is sometimes seen in male bats during their hibernation (317) and in breeding season male short-tailed shrews often show erection during sleep.

In the case of the shrew vaso-dilation of the copulatory organ is sometimes accompanied by additional elements of the coital pattern as indicated in the following quotation. "These [erections] occur while the animal is lying asleep on its side or back. First, the penis becomes erect; then the animal stirs restlessly and there are twitching movements of the hind legs . . . . After several seconds a number of minor thrusts are made, followed by one vigorous, deep thrust which is accompanied by repeated extension of the hind limbs as in actual copulation" (232). This pattern may be repeated two or three times before the animal awakes and several such episodes can occur within half an hour. These observations suggest the interesting thesis that under such conditions sensory impulses originating in the erect penis set off a train of neuro-muscular events reflected in the execution of most of the major elements in the normal mating pattern.

[5] Equally unsatisfactory are theories designating sensory impulses from other parts of the reproductive system as initiating causes. Thus copulation continues in male rats after removal of the vas (12), the seminal vesicles (56), or even the testes providing androgen is administered (266).

The resemblance between such behavior in lower mammals and the nocturnal emissions common in a high percentage of humans[6] is obvious, and although crucial evidence is lacking it may be suggested that the erotic dreams which sometimes accompany genital activity during sleep are initiated or at least reinforced by sensory impulses from the erect penis or clitoris.

6. *Central nervous effects of peripheral stimulation.* Consideration of the observational and experimental evidence indicating that arousal and intensification of sexual excitement normally involves several types of sensory stimulation has led to theoretical analsyis of the more central neuro-physiological processes concerned (41). It has been suggested that effector mechanisms responsible for integration of the motor acts of courtship and coitus not only receive fibers from the several receptor systems, but may also be subject to facilitative impulses arising within a functionally related but structurally independent "central excitatory mechanism" (40). The *c.e.m.* is assumed to maintain the excitability of the executive centers by distant facilitation; and this part of the central nervous system is regarded as the most likely substrate for the effects of multisensory or repetitive peripheral stimulation.

Thus it seems probable that, by virtue of spatial summation, simultaneous stimulation of several sensory receptors may evoke a response within the *c.e.m.* in much the same fashion as the concurrent arrival of two subliminal impulses over different afferents elicits discharge of a reflex center. Furthermore, successive combination may occur within the *c.e.m.* as a result of successive applications of a sub-threshold stimulus to a single effector (128).

B. *Effector Functions.* As explained at an earlier point in this review, the effector or motor functions of the nervous system in sexual behavior are taken to include both the mediation of the separate responses involved in courtship and copulation and the integration of these individual reactions into a well-organized, biologically effective pattern. Most writers agree that the two functions probably are carried out by different nervous mechanisms (187, 213, 265, 331).

1. *Fundamental elements of the coital pattern.* The complete mating pattern is a complex affair including preliminary orientation to the partner, execution of postural adjustments appropriate to coition, and occurrence of the essential genital reflexes.

a. *Sub-primate species.* The basic coital reactions are approximately the same in all sub-primate species (247a). The female's fundamental copulatory response consists of flattening or depressing the back with consequent elevation and exposure of the perineum, and in those species where it is necessary the tail is moved laterally. Primary elements in the male's performance include mounting the female from the rear with the forelegs resting on her back or gripping her sides, execution of piston-like thrusts of the hind quarters and insertion of the erect penis. During coition most male rodents palpate the female's sides with the forelegs. In some species, such as the cat and dog intromission normally occurs but once, and is prolonged until ejaculation takes place. In other species the male may mount, achieve intromission and dismount several times before, on a final copulation, ejaculation occurs.

[6] Reported by 75 per cent of 291 adolescent boys (246).

From time to time there appear in the literature reports of apparent exceptions to this generalized pattern of copulation in lower mammals, and it is amusing to note the consistent anthropocentrism with which observers describe ventro-ventro mating in various species. None of the published accounts bears careful scrutiny, and inaccurate observation or incorrect identification of the behavior observed usually is sufficient to explain the error. For example, reports of ventro-ventro mating in the sea otter seem questionable in view of the fact that the animals under observation were a considerable distance away and were partially submerged at all times (223). Or again, in one study drawings purporting to represent ventro-ventro mating in the hamster are in actuality illustrations of fighting behavior (234). For obvious reasons it is often assumed that porcupines are incapable of mating in the manner characteristic of most quadrupeds, and in fact some naturalists have described the copulation of this species as occurring while male and female stand facing each other. However, careful observation of tame specimens reveals that although pre-copulatory courtship often includes some sparring or wrestling when the pair stand erect (265b), intromission is anatomically impossible in this position. When the actual mating takes place the male approaches the female from the rear, resting his forepaws upon the ventral surface of her tail (which is curved upward and over her back), and the female elevates her perineum to expose the genitalia (265a, 300). The european hedgehog copulates in essentially the same fashion (164a).

The statement that the fundamental responses involved in the copulation of lower mammals appear in relatively stereotyped form should be qualified by the admission that considerable variability is possible,[7] and under experimental conditions individual animals may show a surprising degree of adaptability in their coital reactions. Although they cannot mate in completely normal fashion bipedal male and female rats from which the front legs have been removed in infancy are capable of fertile copulation with normal individuals of the opposite sex (242).

b. *Primates.* The copulatory pattern common to most sub-human primates resembles in many respects that of the lower mammals, although differences in body form are of course reflected in behavior. In coition the receptive female rhesus monkey bends forward at the hips, lowering the head and shoulders and presenting the genitalia; while the male rests his hands upon the female's hips and grips her hind legs with his feet. Two to eight pelvic thrusts occur in a single intromission and several mountings usually are necessary before ejaculation occurs (75).

The copulatory pattern characteristic of the adult chimpanzee is closely similar to that of lower primates, but detailed descriptions of pre-puberal sex play in this anthropoid species include repeated observations of the occurrence of ventro-ventro coition. This variant of the pattern usually begins when the male and

[7] Under certain conditions the male may attempt to mount the female from almost any angle, and highly excited male rats aroused by previous copulations with an estrous female may clasp and palpate a non-receptive female even while she is lying on her back in a defensive position, a reaction which the naive observer might mistake for true copulation.

female are seated embracing each other in a face-to-face position. The female leans backward drawing the male over her, her legs encircling his back. The male's feet are flat on the floor, and his upper legs embrace the female's pelvis, hips and thighs. Pelvic movements follow and intromission occurs in some instances (55).[8]

2. *Effector mechanisms in the spinal cord.* a. *In the female.* Evidence to be reported immediately below concerning the spinal integration of genital reflexes in male animals makes it appear probable that the female's clitoral reactions are mediated by neural circuits in the cord. There is no direct experimental evidence to support this hypothesis but it would seem to be in harmony with the fact that these simple genital responses appear very early in the post-natal life of female anthropoids. Thus, erection of the clitoris occurs in the female chimpanzee within the first few days after birth (55), and is present as early in the female human infant (208). Certain clinical observations also suggest the possibility of "partial orgasm" in the infant human, for it is reported that manipulation of the turgescent clitoris is followed by "unmistakable relief from tension and restlessness" (208).

Despite occasional contradictory reports it is fairly well established that neural elements within the spinal cord are capable of supporting many of the basic responses which constitute the adult female's copulatory pattern. Although female guinea pigs in which the cord is sectioned at any point between T5 and L2 are reported to show no estrous behavior (94), the bitch exhibits normal estrous and mating despite complete lumbar section (136); and female cats with the cord transected at C1 display treading, opisthotonus and lateral deviation of the tail in response to stimulation of the perineum. Some investigators have reported that the spinal cat shows this behavior only during natural estrus or under the influence of injected estrogen, and these workers conclude that the responses in question are short-arc reflexes which in the intact animal are phasically integrated by higher centers (197, 198).

There is some dispute regarding the genuineness of the estrogenic effect in this instance, and other writers claim that chronic spinal cats exhibit no "estrous" reactions (24). There is no attempt to deny the appearance of the responses described above, but it is stated that lateral deviation of the tail is shown by males as well as females, that treading and opisthotonus may be elicited in the anestrous

[8] It is of general interest to compare coital performance in the ape with ethnological descriptions of the copulatory techniques employed in various human cultures. In a few tribes the position most commonly adopted during intercourse resembles closely that of the other anthropoids, the male approaching the stooping female from the rear (122). More wide spread is a variation of the ventro-ventro position comparable to the foregoing description of prepuberal copulation in the chimpanzee. For instance, copulation as practiced by all tribes among the Australian Aborigines of the North-West-Central Queensland district has been described as follows: "The female lies on her back on the ground, while the male . . . sits on his heels close in front: he now pulls her towards him and raising her buttocks drags them into the inner aspects of his own thighs, her legs clutching him around the flanks . . ." (256). This general pattern of coitus is said to be the common form in most primitive societies, whereas the ventro-ventro position employed in many European cultures is regarded by primitives as less effective and even "abnormal" (122).

female, and that since they are not dependent upon estrogen such reactions cannot be termed "estrual" (24).

At this point in our survey we need not be concerned with the presence or absence of hormonal effects and it is sufficient to note that treading, opisthotonus and tail deviation, which are components of the normal feminine coital pattern, survive in the spinal preparation.

Female cats decerebrated at the pontile level respond to mechanical stimulation of the vulva with collapse of extensor rigidity of the forelimbs, and the resemblance of the resulting posture to the sex crouch has suggested to some observers the existence of a primitive bulbo-spinal mechanism mediating this element in the female's copulatory pattern (23).

b. *In the male.* That certain elements of the male mammal's copulatory pattern can be mediated by spinal mechanisms is shown by the elicitation of sexual responses upon stimulation of the cord, and by the survival of such reactions after cord section.

Application of a low voltage current to the sacral region of the cord evokes erection and ejaculation in male rabbits and rats (101), cats (264) and dogs (52), and trauma attendant upon cord section elicits the same responses in male guinea pigs (172). In guinea pigs transection between T12 and L1 is followed within one to seven minutes by "rhythmic movements of the ano-genital region,"[9] erection and ejaculation (8). The common appearance of priapism in hanged criminals (128) is probably attributable to traumatic stimulation of similar spinal mechanisms.

Evidence for survival of certain elements in the masculine sexual pattern in chronic spinal animals is to be found in the reports of several investigators. The spinal male cat exhibits erection when placed in a prone position which brings the genitals into contact with the substratum (102). Following supra-lumbar transection male dogs show erection and ejaculation in response to genital manipulation (23), and these reactions can also be elicited in male humans surviving the same operation (24).

Sexual activities integrated within the cord are not limited to genital responses, for certain gross postural adjustments characteristic of copulation may also be mediated at this level. After lumbar section male dogs respond to genital manipulation with a series of reactions that have much in common with the normal copulatory pattern. These include bilateral extension at the knees, ankles and hip joints, depression of the tail and downward curving of the pelvic portions of the body (261). Masturbatory stimulation of spinal man evokes a composite response (of which of course the patient is quite unaware) which is so obviously similar to copulatory activity that it has been termed the "coital reflex" (251a).

Descriptions of the survival of isolated sexual responses in spinal preparations should not be allowed to obscure the fact that these reactions are disconnected fragments of a complete pattern, and that in the intact animal the functional

[9] Although not specifically described, these movements are suggestive of copulatory thrusts.

activity of the spinal mechanisms undoubtedly is subject to extensive modification by impulses from higher regions of the nervous system.[10]

3. *Effector mechanisms in the sympathetic system.* Complete abdominal sympathectomy does not prevent fertile mating in the female rabbit (65), rat (9) or cat (21); but sympathetic innervation appears to be heavily involved in at least one primary element of the male's copulatory pattern, namely, the ejaculatory response.

Male guinea pigs subjected to abdominal sympathectomy or to section of the hypogastric nerves respond to subsequent transection of the spinal cord with the display of erection, but ejaculation does not occur (8). In this same species electrical stimulation of the sympathetic pre- or post-ganglionic fibers serving the genital region evokes ejaculation without erection (8); and in the male cat stimulation of the internal pudendal nerves elicits the ejaculatory reflex (264).

Failure of ejaculation due to interference with the sympathetic supply to the genitalia produces marked abnormalities in the copulatory behavior of male animals. Guinea pigs suffering sympathetic denervation of the genitals copulate repeatedly but do not ejaculate (10); and after the same operation male rabbits mate vigorously, but the responses accompanying ejaculation (utterance of a characteristic cry and falling off the female) are lacking (306). In the rabbit inability to ejaculate is reflected in persistent copulatory acts which continue unabated until both the male and female are totally exhausted (8). In the case of the male rat abdominal sympathectomy eliminates ejaculation (306) but does not abolish copulation; and sympathectomized rats have been described as "infertile" (8) or "impotent" (172) although the latter term is of questionable applicability inasmuch as active copulation survives.

Male cats in which the lower two lumbar segments and all of the sacral cord have been removed seize the estrous female, exhibit erection and copulate normally. However, if the operation includes extirpation of the abdominal sympathetic chains or removal of the inferior mesenteric ganglia, penile erection does not occur although the male still grips the female and attempts to copulate (255).

4. *Effector mechanisms in the diencephalon.* a. *In the female.* Nervous mechanisms lying within the diencephalon contribute to integration of the separate responses which comprise the copulatory pattern of female rodents, lagomorphs and carnivores. Results of several operative experiments indicate that in the case of the female guinea pig and rat the most important neural circuits are to be found within a restricted region of the hypothalamus. Mating behavior persists in both species despite transection of the brain just anterior to the mammillary bodies, but is eliminated by section posterior to the mammillary bodies and anterior to the superior colliculus (94). According to other workers the

[10] Direct physiological evidence for functional interrelationships between spinal and supra-spinal mechanisms is seen in the report that penile erection is difficult to elicit in the decerebrate male cat (23). Furthermore, in the intact male dog erection or detumescence can be conditioned like any other reflex, and the appropriate genital response can subsequently be evoked by application of the conditioned stimulus, e.g., the sound of a bell (207).

guinea pig's receptive responses depend upon only the most ventral portions of the anterior hypothalamus, and the mesencephalic tegmentum or the region of the mammillary bodies is not involved (117).

Destruction of the caudal hypothalamus is reported to have no effect upon mating in the female cat, whereas extensive invasion of the anterior hypothalamus eliminates estrous behavior in some females and fails to do so in others (24). In this case no essential mechanism is indicated, but these results suggest that certain types of hypothalamic injury may reduce the female's tendency to display mating reactions.

Although lesions interrupting the optico-hypophyseal tract are reported to eliminate estrus and copulation in the female cat (117), the behavioral change probably reflects interference with secretory function of the pituitary (with consequent ovarian failure and disruption of gonadal hormone production) rather than destruction of any critical executive centers. Descriptions of amenorrhea and loss of libido following injury to the floor of the third ventricle in the human female (301) are perhaps susceptible to similar interpretation.

b. *In the male.* In view of the apparent importance of this brain region to the mating activities of female mammals it is unfortunate that few experiments have dealt with the effect of diencephalic injury upon sexual behavior in the male. The single report available states that damage to the medial half of the hypothalamus tends to "depress" sexual behavior in male rats, although the general area apparently includes no structure essential for the integration of the copulatory pattern (78).

It has been observed that injury to the floor of the third ventricle may abolish sexual desire and potency in the human male (301), and a study of thirty-three men with head injuries in which the diencephalon was the focus of damage revealed complete or partial loss of sexual desire in the majority of cases (285). However, evaluation of such reports is especially difficult in view of the possibility of direct or indirect interference with pituitary function, with resultant degeneration of the reproductive glands and consequent elimination of testicular hormone.

5. *Effector mechanisms in the forebrain.* a. *In female rodents, lagomorphs and carnivores.* Observations to the effect that copulatory responses may survive in female mammals despite complete transection of the brain at the level of the anterior hypothalamus indicate that more rostral parts of the brain are not essential to such reactions, and results of various experiments confirm this view.

Injury to or complete removal of the cerebral cortex, hippocampus, and the caudate-putamen complex does not abolish coital activity in the female guinea pig (94), rabbit (65), rat (92) or cat (22). There is in fact some indication that neo-cortical injury may increase the frequency of certain mating reactions. Completely decorticated female rats exhibit an increased tendency to assume the copulatory position, display ear-vibration more frequently and show a decrease in the frequency of the back-kicking response by which the intact female occasionally repulses the male (45). Bilateral temporal lobectomy in female rhesus monkeys is followed by increased willingness to receive the male and by heightened frequency of masturbation (177).

Evidence for the survival of certain estrous reactions in partially or completely decorticated females should not be taken to mean that the mating behavior of the intact animal occurs independently of activity within the neo-pallium, for it is clear that the cortex normally plays an important part in the mediation of feminine receptivity. One element in the female rat's mating pattern consists of a tendency to seek out the male and to orient her pre-copulatory heat responses toward him (313). The decorticate female of this species shows frequent estrous reactions but exhibits no tendency to direct such behavior toward the male. Furthermore, although all elements of the normal copulatory pattern may survive total loss of the neo-pallium, the timing or integration of the separate responses usually is disrupted (45).

b. *In male rodents and lagomorphs.* There are no published reports to indicate the effects of exclusively non-cortical forebrain injury upon mating activity in male animals, for all investigations have dealt with the results of decerebration or with the effects of lesions restricted to the neo-cortex.[11]

It is known that grossly normal copulatory reactions survive in male rabbits despite removal of all neo-cortical tissue (65), although no quantitative data have been presented to indicate possible changes in sexual responsiveness after this operation. Mating reactions are also reported to have been shown by four of eight decorticate male rats (92), but in an earlier publication I have questioned the validity of these conclusions on the basis of the author's failure to reconstruct cortical lesions (which were probably much smaller than estimated) and because no direct observations of mating behavior were conducted (31).

More systematic studies of pre- and post-operative sexual behavior in cortically injured male rats have revealed several significant facts. Lesions destroying no more than 20 per cent of the cerebral cortex do not eliminate mating behavior regardless of the region or regions involved (31). Such findings are in opposition to earlier proposals concerning the existence of a focalized "sex center" in the cortex (77).

Destruction of 20 to 50 per cent of the cortex may or may not eliminate copulatory behavior (184, 292). Individual differences are pronounced, but if lesion size is held constant within these limits, coital activity tends to survive in males which were sexually very vigorous before operation and to be eliminated in those animals which showed a low pre-operative sex drive (46). As the amount of neo-cortical destruction increases above approximately 20 per cent there occurs a progressive decrease in the proportion of males continuing to copulate; and when the extent of bilateral invasion exceeds 60 to 75 per cent of the total neo-pallial mass sexual reactions are eliminated in all cases (32).

Quantitative methods for measuring and describing mating activity in male rats yield results which emphasize the inadequacy of any attempt to interpret the behavioral effects of neural injury merely in terms of presence or absence of a particular response. For example, even in those males which continue to mate despite cortical injury, the tendency to display sexual reactions decreases pro-

[11] I have observed normal copulation in a few male rats suffering destruction of approximately one-fifth of the caudate-putamen complex without cortical injury, but no quantitative tests were conducted and the findings have not been published.

portionately with increase in lesion size. Thus, although some individuals may continue to copulate after loss of half of the neo-cortex they mate in only 15 per cent of the post-operative sex tests, whereas coition had occurred in 100 per cent of the tests given before operation (31).

The exact nature of the cortical contribution to sexual behavior in the male rat is not understood, but it seems safe to conclude that the neo-pallium is not essential to the control of the motor pattern involved, for when brain-operated individuals copulate the overt reactions are grossly normal (92). Studies of the effects of restricted cortical injury upon sensory capacities in rodents have yielded results which suggest that absence of sexual behavior in decorticate animals cannot be explained as the result of intereference with major sensory pathways or areas of representation (46).

The behavior of decorticated males placed with females in heat is characterized not by abortive attempts to copulate nor by any other sign of interference with the overt expression of sexual excitement, but by the absence of any indication of primary arousal. Neo-cortical injury appears to decrease sexual responsiveness without seriously affecting either the capacity for reception of stimuli or for execution of the motor acts of courtship and copulation.[12]

Many isolated masses of nerve cells in the central nervous system appear to be concerned with regulation of general levels of excitement and to contribute little if anything to the specific patterning of reactions (186). For example, studies of the function of the bulbo-spinal system in animal locomotion have led to the proposal that "one may look upon the medulla as a *Stimulations organ* which maintains the excitability of the spinal cord at a level which enables the latter to display its full locomotor activity when exposed to extraneous stimuli at a comparatively low level of intensity" (143).

In the case of the male rat it is apparent that nervous mechanisms lying below the neo-cortex are responsible for organization of the overt acts of courtship and copulation, but the reactivity of the executive circuits to external stimulation depends in part upon some function of the cerebral cortex.[13] The specific nature of the cortical influence is not known, but it probably rests upon well-known neuro-physiological phenomena such as that of "distant facilitation" (69).

Relations between the cortical mechanisms and lower nervous circuits involved in sexual reactions should be conceived as reciprocal rather than unilateral, for

[12] This reduction in reactivity is, of course, not restricted to the sexual field. On the contrary, it is a widespread effect which influences many types of behavior; and concepts of decreased "exploratory activity" (185), reduced "vigilance" (160), lowered "attention" (187), altered "affective regulation" (236) and increased "apathy" (114) reflect the general recognition of decreased responsiveness consequent to brain injury.

[13] As will be shown later the excitability of the executive circuits is also altered by gonadal hormones. Normal sexual responsiveness in the male depends not only upon neural impulses from the cerebral cortex but upon chemical sensitization by testicular hormone. In intact animals administration of large amounts of androgen may raise the reactivity of the copulatory mechanisms well above normal (34); and in partially decorticated cases mating responses can sometimes be revived by androgen treatment (31, 46). It appears that in the latter instance the hyper-normal hormone level exerts a compensatory effect, maintaining normal reactivity in the integrative mechanisms for sexual behavior despite reduction in the facilitative influences normally provided by the cortex.

impulses of diencephalic, mesencephalic or even bulbo-spinal origin probably are capable of modifying the sexual functions of the neo-pallium. It is in fact known that, in cats, stimulation of specific hypothalamic nuclei facilitates certain types of activity within the cortex, and it has been suggested that "initiation or accentuation of motor phenomena during emotional states may have a basis in hypothalamic-cortical facilitation" (224).

The data presented above indicate that sexual arousal and mating responses are to be considered as manifestations of an extremely complex pattern of neurophysiological activity involving mediation of some response items by spinal centers, phasic integration of separate reactions by higher circuits located between the cord and cerebral hemispheres, and facilitative or inhibitory interaction between the foregoing mechanisms and the cerebral cortex.[14]

In male carnivores. It is usually accepted that the specificity of cortical influence upon sensory-motor functions in different species is correlated with the general level of development of the neo-pallium. In contrast to smooth-brained rodents, carnivores possess a convoluted and fairly well-differentiated cerebral cortex, and it is therefore not surprising to learn that male cats sustaining particular types of cortical injury show disturbances of sensory-motor co-ordination which are reflected in the occurrence of inaccurate and abortive mating reactions.

Coital responses are reported to have survived in six of seven male cats suffering bilateral removal of the frontal cortex in combination with all tissue anterior to this region; but sensory-motor abnormalities consequent to the operation resulted in a high incidence of disoriented reactions to the female (183). An increase in sex drive is said to have occurred in some of the brain-operated males, but it is impossible to assess the significance of this statement in the absence of supporting quantitative data (e.g., decreased latency or increased frequency of copulation, more frequent or rapid attainment of ejaculation, etc.).

In an experiment which has not yet been published Dr. Arthur Zitrin and I tested the mating activity of twenty male cats before and after destruction of various parts and amounts of the cerebral cortex. We found that normal copulation was possible after unilateral hemi-decortication, or bilateral removal of the frontal, temporal, parietal or occipital lobes. However, although they mated normally in some post-operative tests, at other times animals with such lesions exhibited disoriented or unco-ordinated sexual responses. For example, operated males sometimes approached and gripped the female from the side instead of the rear, and then displayed erection and pelvic thrusts although they were not mounted and the genitalia were not in contact with any portion of the female's body.

Cats subjected to unilateral hemi-decortication in combination with destruction of the contra-lateral frontal pole exhibited a strong sex drive as reflected in re-

[14] The concept of cortical function in sexual behavior presented here resembles in several ways interpretations advanced more than thirty years ago in an attempt to combat contemporaneous theories of a cerebral sexual center. At that time it was suggested that the rôle of the cortex in mating, as in most vegetative and visceral activities, is generalized rather than specific; although it was recognized that in higher forms the neo-pallium may play an important rôle in the associative learning which directs and coordinates specific sexual reactions (221).

peated attempts to copulate and in the frequent occurrence of erection; but these animals were capable of coition only when they were placed upon the receptive female. Hemi-decortication with destruction of the contra-lateral occipital pole eliminated responses to the distant female; but males sustaining lesions of this type were capable of normal orientation and copulatory reactions as soon as they came into physical contact with the estrous queen.

Cats in which all of the neo-cortex had been removed showed no interest in the receptive female and made no attempt to copulate even when placed upon the stimulus animal. Other workers have observed that decerebrate cats and dogs show no mating reactions (24).

In such cases it is difficult to determine the extent to which sexual failure is due to interference with primary sensory capacities consequent to removal of cortical projection areas. It is known that destruction of the striate area in dogs and cats does not abolish all visual function (131), and in the latter species this operation produces no marked defect in optical acuity as measured by nystagmus induction (271). Furthermore, dogs subjected to extirpation of the auditory cortex retain the ability to discriminate sounds (2), and the acuity of hearing is only slightly impaired (132). Some sound discrimination is present in cats after removal of the entire neo-pallium (304). Bilateral loss of parietal, temporal, frontal or occipital lobes does not appear to depress olfactory or gustatory functions in the dog (1).

In spite of this sort of evidence, which indicates partial survival of sensory capacities in feline and canine species after destruction of various cortical projection areas, it is beyond question that extensive neo-pallial injury produces marked reduction in sensitivity to external stimulation, and one is obligated to consider the possibility that absence of copulatory reactions in decorticate or decerebrate male cats and dogs is due solely to interference with exteroceptive function.

Until more data are collected it will be impossible to prove or disprove any such interpretation; but it should be noted that the sensory impairment consequent to bilateral destruction of occipital, temporal, parietal or frontal lobes does not prevent copulation in male cats. Furthermore, even the pronounced sensory-motor disturbance produced by unilateral hemi-decortication combined with removal of the remaining frontal pole fails to obviate sexual arousal and coital attempts. After complete decortication, however, males give no sign of sexual arousal and make no response to the receptive female. When these cats are placed upon the female in the mating position they show neither erection (which is reported to occur in the spinal cat) nor pelvic movements. In harmony with these observations are the reports of other workers to the effect that decerebrate male dogs, "show no interest in the rutting bitch" (138).

It seems clear that in the case of male carnivores the cortex makes a more specific contribution to sexual performance than it does in rodents. The question remains as to whether, in addition to particular sensory-motor functions, the carnivore cortex also exerts a general facilitative influence upon lower centers more intimately involved in the mediation of courtship and copulation. At

present all that can be said is that such an effect is quite possible; for the striate cortex is known to exert non-visual as well as visual effects (188), and studies of the primate brain have led to the conclusion that several cortical projection areas, "participate in a more general type of integration that is dissociated from their functions as sensory regions" (173).

d. *In male and femalr primates.* There are no reports dealing with the effects of extensive cortical injury upon sexual activity in male or female primates. In connection with investigation of the psychological changes accompanying brain injury in the rhesus monkey it was noted incidentally that males suffering bilateral temporal lobectomy exhibited a striking increase in sexual drive. The important behavioral alterations in males included priapism combined with a marked rise in manual and oral masturbation and in the frequency of homo- and hetero-sexual intercourse (177); and females with similar lesions are described as giving similar indications of hypernormal sexual responsiveness. The significance of this report is not immediately apparent, but it will be recalled that loss of the frontal lobes is said to have produced increase in the sexual responsiveness of male cats (183).

Existing knowledge regarding the extensive encephalization of various functions in the primate brain (164) renders it highly probable that at this level of the phyletic scale removal of sensory and motor projection areas would produce marked abnormalities in the mating pattern of both sexes and might abolish all responsiveness on the part of the male.

6. *Evolutionary trends.* Comparison of the effects of forebrain injury upon the reproductive life of fish (227, 229), amphibia (6), reptiles (97), birds (254) and mammals indicates that this most recently evolved part of the central nervous system has come to play an increasingly important rôle in the control of courtship and mating responses (42, 227). The most obvious type of change at ascending levels within the mammalian order lies in the growing influence of cortical areas of representation involved in sensory and motor activities Progressive encephalization of sensory-motor functions has naturally affected all complex types of behavior, and this process seems to carry with it certain secondary alterations which are of particular importance in connection with the response patterns under consideration here.

It appears that as the lower neural mechanisms mediating various discrete sexual reactions come to be more and more dependent upon activity within the neo-cortex, several associated changes occur. (1) The variety of external stimuli capable of evoking sexual responses tends to increase; (2) the overt patterns of sexual expression tend to become more variable; and (3) the direct importance of the gonadal hormones to sexual behavior tends to decrease. These generalizations, which are based upon scattered and fragmentary evidence, undoubtedly are subject to many exceptions, and may eventually prove to be only poor approximations of the facts; but at the present state of our knowledge they provide a useful frame of reference which should be helpful to an analysis of the behavioral and endocrinological data that follow.

C. *Behavioral Data Related to the Neuro-Physiological Evidence.* Having sur-

veyed the physiological evidence concerning neural mechanisms responsible for some aspects of courtship and mating we may now profitably turn to the consideration of a complementary body of data consisting chiefly of observational or descriptive accounts of sexual activity in different mammalian species. Some of the material presented below offers support for conclusions based upon neurophysiological experiments. In other instances the behavioral evidence extends beyond results established by physiological techniques and points the way toward potentially fruitful laboratory studies.

1. *Relation between sexual behavior and non-sexual emotions.* Little is known concerning the neural basis for emotions in general, but observations of behavior strongly suggest a close functional relationship between the mechanisms for sexual behavior and other nervous structures which are activated during various sorts of intense emotional arousal. It is particularly evident that sexual responses may appear in non-sexual environmental settings which evoke strong emotions of different types; and conversely, stimuli which ordinarily elicit courtship and copulation may under certain conditions call forth vigorous non-sexual reactions indicative of high excitement.

That sexual mechanisms may be activated by non-sexual stimulation is shown by the fact that male dogs sometimes display erection and ejaculation in the presence of fear-inducing stimuli (129), and that the same responses frequently occur in male chimpanzees subjected to various types of excitement such as that engendered by presentation of food, separation from companions, or disciplinary contact with the trainer (55). Directly comparable evidence is available in the case of humans, for painful stimuli may evoke genital reflexes in young children (171), and pre-adolescent boys experience erection when frightened, as well as in other exciting but non-sexual situations (246). Sexual presentation is sometimes exhibited by both female and male rhesus monkeys in response to "generalized, non-sexual excitement" (75).

The sexual reactions which appear in conjunction with intense emotion are not limited to fragmentary, genital reflexes. On the contrary active sexual behavior may occur in such situations. For example, immature chimpanzees occasionally masturbate during a fit of anger (178), and cebus monkeys as well as apes sometimes indulge in manual or oral genital manipulation when confronted with a difficult learning problem (129). Similarly it is reported that school children often resort to genital play under the stressful influence of classroom examinations (171, 172). In human adults extreme frustration may lead not only to auto-genital activity but to complete hetero-sexual intercourse (260).

A state of generalized excitement seems in some cases to facilitate the occurrence and manifestation of sexual arousal. Female rats which have escaped from their cage and been recaught sometimes exhibit an increase in sexual excitability (16); and some males of this species which have become temporarily sluggish can be aroused to renewed copulatory activity if they are batted sharply about the cage by the experimenter (40). In some instances male rats which fail to initiate coital activity in response to a given estrous female may do so when a new stimulus animal is introduced, even though the second female is no more

receptive than the first (298). In this case it seems probable that the disturbance created by changing females is responsible for the male's arousal; an interpretation which gains support from the observation that the same result may be obtained by removing the original female and immediately re-introducing her into the testing cage (34).

Male rats which are regularly subjected to an electric shock just before and during the moment that a receptive female is placed in the cage react in subsequent tests by mounting the female the moment she is presented, and "copulating violently, almost frantically, *during the shock*" (314). This reaction is the opposite of that obtained if the shock is started when the copulation is in progress, or just after it is terminated. In the latter cases punishment inhibits sexual activity in following tests.

Male and female primates are known to show sudden increases in mating activity in response to exciting but non-sexual changes in the environment; and in some instances the change may be nothing more than the abrupt appearance of a human observer (55, 175). It is quite possible that the stimulating effect of non-sexual excitement is responsible for the discriminatory behavior shown by males of various species toward new or strange females. In free-ranging rhesus monkeys the stimulus value of a new female is greater than that of a female with which the male has been associated for some time, even though the latter animal may be fully receptive (75, 76). The result is that the male displays an increase in copulatory behavior when a new female becomes available. Human parallels are too obvious to deserve comment.

A high degree of sexual arousal often produces non-sexual manifestations of excitement, and this is particularly common when normal coital adjustments are blocked. For example, male rats with under-developed penes execute repeated attempts to copulate with the receptive female, but in the absence of intromission ejaculation is rare. When such animals become highly excited as a result of a series of incomplete copulations the frequency of attacks upon the female is much higher than in normal matings (50); and this increased aggression seems to be related to the occurrence of intense sexual arousal combined with the failure of ejaculation and orgasm. Female rats which are extremely receptive often launch vicious attacks upon males which persist in investigating and mounting but fail to copulate and achieve intromission.

The literature contains one report of an inexperienced male macaque which became intensely excited as a result of his inability to achieve intromission even though the female was highly receptive and attempted to guide the penis manually. After several unsuccessful attempts at coition the male left the female and began to bite his feet and legs, inflicting severe wounds. Thereafter the mere appearance of the female tended to throw this male into a frenzy of self-mutilation (310).

2. *Inter-sexual differences in sub-primate species.* Behavioral evidence which correlates with the results of certain physiological investigations is found in a list of differences between the sexual behavior of males and females of several lower mammalian species.

a. *Effects of non-sexual stimuli upon sexual arousal.* Among lower mammals extraneous, non-sexual stimuli have a tendency to inhibit sexual arousal in the male. Males of many species are inattentive to receptive females and slow to copulate in surroundings that are novel and strange; whereas under the same circumstances females usually mate without hesitation. This inter-sexual difference is recognized by commercial breeders (5) and on mink ranches, for example, it is customary to take the estrous female to the male's living quarters for service; for when the reverse procedure is practiced the male, "wastes time exploring the cage, etc." (179).

The absence of sexual arousal in a strange environment may be due to fear, but more frequently it appears that the male is merely distracted by non-sexual stimuli, a situation which occurs less commonly in the female. Thus although captive female coyotes permit copulation by male dogs, the reciprocal cross is difficult to bring about because the male coyote is easily upset by the artificial conditions of caging, etc. (95).

The sexually inhibiting effects exerted by a strange environment may be dissipated gradually as the male becomes accustomed to his new surroundings, and some individuals which fail to mate at the first contact succeed in doing so after they have been put with a receptive female in the same cage several days in succession. In repeated tests the proportion of male rats exhibiting coital reactions tends to increase (36); a result which is not typical of females and which presumably reflects individual variance in the readiness with which different males adapt to the testing situation.

In other cases environmental interference is promptly and completely eliminated as the result of one successful copulation. Regardless of the duration of their residence in the laboratory most male cats do not copulate when they are first taken to the testing room and placed with a receptive queen, and this refractory behavior tends to continue indefinitely until coition is once achieved. In some instances it is necessary to manipulate the experimental situation in order to induce the initial mating; but once this has taken place the male thenceforth displays spontaneous courtship and copulation in all subsequent periods (334). Estrous female cats, in contrast, mate readily during the first test under laboratory conditions.[15]

In some species the male is hesitant to mate with a strange female even though she is fully receptive, but comparable discrimination is not evidenced by the female. For example, the receptive female porcupine accepts any sexually aggressive male upon first contact, whereas males will not attempt coition with

[15] The observation that females are less likely than males to be distracted by non-sexual environmental stimuli finds its counterpart in the fact that females of several lower mammalian species are fully capable of attending simultaneously to sexual and non-sexual stimuli; whereas the male gives no indication of this ability. The estrous bitch accepts food even during coition, while the male is capricious about eating in the presence of a receptive female and most males refuse food entirely during copulation (129). I have seen female cats *in copulo* conduct extensive olfactory investigation of a mouse hole in the board floor of the testing room, although such a division of activity has never been observed in the case of the male.

any female until she has been kept in their cage for several days (265a). Males of various domesticated species sometimes refuse to serve certain types of females. A stallion which has mated exclusively with dark mares may show no inclination to mount a white one; and jacks which have been allowed to serve jennets are likely to refuse copulation with mares (206a).

The converse of environmental inhibition of sexual responsiveness can be seen in the behavior of males of several lower mammalian species. When a male has repeatedly received receptive females in a given external situation, the environmental setting tends to acquire sexually-exciting qualities with the result that coital reactions may be elicited by any one of a wide variety of stimulus animals offered in that setting. Although sexually-inexperienced male rats usually attempt to copulate only when given a female rat in estrus (41), many males which have copulated with receptive females in a particular testing cage will subsequently attempt to mate with males or non-receptive females (30, 174, 286), or even with females of a foreign species (41) when such animals are encountered in the original testing situation. Male rabbits accustomed to receiving estrous does in the living cage will, "mount cats, kittens, guinea pigs, and even dead or inanimate objects" offered in a similar fashion (65).

The sexual responsiveness of male mammals may be depressed in an environmental setting associated with conflict or pain. Neurotic male dogs are slow to mate with estrous females presented in the experimental room where the neurosis was induced, although in the paddock copulation may occur promptly (129).[16] Lasting inhibition of the male's sexual responsiveness sometimes results if arousal is attended by punishment. Male rats subjected to electric shock during or immediately after copulation with the female usually fail to mate in subsequent tests (314). Male marten which are scratched or cut by the female during mating are apt to show little interest in receptive females presented on succeeding days (204).

Experiences encountered during the first complete sexual contact exert a particularly potent effect upon subsequent responsiveness. In training young male foxes for service, breeders take care to select females that are gentle and in full heat in order that the male shall not be attacked or frightened and thus spoiled for future use as a sire; but such precautions are regarded as superfluous in the case of an experienced stud male (108). If a fight occurs during initial matings the young male mink is unlikely to develop sufficient aggressiveness toward the female; whereas after a few successful copulations, "he will have acquired confidence and his future as a breeder is assured" (179).

The facts that extraneous stimuli tend to interfere with the male's reactivity to the receptive female, and that previous experience may alter the male's susceptibility to sexual arousal appear to be closely related to the observation

[16] Evidence of this type may bear a significant relationship to clinical observations suggesting that conflict, anger or fear are capable of suppressing sexual tendencies in the human (74); that the majority of patients suffering "anxiety attacks" complain of sexual difficulties (129); and that interference with normal sexual function is a common symptom in various psychoses, particularly paranoia (306).

that the cerebral cortex exerts a profound effect upon copulatory behavior in the male but is less important in the coital adjustments of the female. In the case of the male a profusion of non-sexual stimuli may evoke within the cortex a series of activities incompatible with those essential to sexual responsiveness, with the result that the effector mechanisms for courtship and copulation receive no facilitative impulses from higher centers. Conversely, stimuli which originally possess little or no effect upon sex responses may, by a process of cortical conditioning, come to exert pronounced effects upon sexual mechanisms.

b. *Pre-puberal sexual activities.* In many sub-primate species there is a marked inter-sexual difference in the age at which recognizable copulatory reactions appear. Although female rats do not exhibit any part of the adult feminine coital pattern prior to the first complete ovarian cycle (which usually occurs at 45 to 50 days post-partum), males of this species mount and palpate other rats and may show pelvic thrusts as early as twenty-one days after birth (43). Comparable pre-puberal sexual responses are shown by male guinea pigs (193), hamsters (62), marten (204), cattle (169) and lions (88); but in none of these forms does the pre-pubescent female display the complementary feminine reactions.

These extreme differences probably are not due exclusively to neural factors, but are related in part to differential rates of development within the secretory portions of the male and female gonad. Thus the interstitial tissue of the testis matures earlier than its ovarian counterpart (27, 311), and produces hormone long before the ovary begins to elaborate and release its endocrine products (27, 169, 231). It seems likely, however, that the relatively early appearance of masculine sexual performance and the delayed onset of feminine receptivity is to be understood as depending upon central nervous differences as well as hormonal variation; for it has been shown that adult male rats which were castrated on the day of birth exhibit sexual responses that are identical to those characteristic of the intact, immature male (50).

3. *Sexual differences and similarities in primates.* Male and female monkeys, apes and humans do not exhibit the same inter-sexual differences as have been described for lower mammals, but certain types of variation which are apparent suggest that at these higher evolutionary levels the nervous elements involved in sexual arousal and coital performance are not identical in the two sexes.

a. *Pre-puberal sexual activities.* Both male and female primates display sexual responses during early life. Penile erection is present in the macaque soon after birth, and by the seventeenth post-natal day males repeatedly rub the glans against objects in the environment. Rhythmic sex patterns become a regular habit during the first year, and male infants use rubber playthings "in a sexual fashion" (166). The female rhesus shows sexual presentation as early as six months after birth. Infant chimpanzees of both sexes indulge in frequent and highly variable sex play which ranges from manual and oral masturbation to heterosexual copulation with intromission and pelvic thrusts (55).

In view of the common occurrence of pre-puberal sexual activities in every primate species which has been investigated it is not surprising that sex interests

are detectable in very young children of all cultures (122). In some social settings infantile sexual tendencies may not become well patterned until puberty (139), but social taboos against overt sexual relationships do not completely suppress such activity (245, 263).

From the evolutionary point of view it is especially significant that in cultures which permit free and unrestricted sexual expression, the common patterns of pre-puberal play bear close resemblance to those observed in the immature chimpanzee, including genital manipulation, oral contact and copulatory attempts with intromission and pelvic thrusts (201).

b. *Sex differences in adulthood.* Published accounts of sexual behavior in primates do not contain as many indications of inter-sexual difference as were cited above in connection with sub-primate species, but studies of reproduction in the chimpanzee indicate that certain important male-female variations do exist.

Adult male chimpanzees lacking experience in heterosexual relations often are unable to effect complete union on first contact with the receptive female. Sexual arousal appears to occur normally, for erection takes place promptly and the male usually attempts to cover the female. In most cases, however, the pattern goes no farther than this, and intromission and ejaculation are not achieved. In subsequent matings the copulatory pattern is gradually perfected and eventually sexual competence is achieved as a result of practice and learning (326).

The female chimpanzee lacking any mating experience is obviously better prepared for copulatory behavior than is the inexperienced male (55). When inexperienced females are caged with experienced consorts mating is carried out without undue difficulty, and even when the partner is a sexually naive male, the female's responses are adequate and may even include manual attempts to direct the penis into the vagina (326).

It is impossible to assess the importance of training and experience to sexual performance in the human. In the first place too little is known concerning the amount and diversity of sexual experience gained before reproductive maturity by members of either sex within our own culture. It seems fairly certain that the overt pre-marital sexual experience of the average male far exceeds that of most females (176); but the breadth of vicarious experience made possible by language would appear to obviate total naiveté in any individual.

The original data have been subject to a great deal of just criticism, but it may be worth while to point out that descriptions of "feral" man indicate a very low sex drive rather than an intense and undisciplined libido. The adult males Tomko of Zips and Kaspar Hauser are reported to have shown no interest whatsoever in sexual matters; and careful consideration of similar cases has led one reviewer to conclude that as far as humans are concerned, "the social channeling of sex is much more necessary for its full expression than might be thought" (333).

4. *Bisexual mating reactions.* Experimentalists ordinarily think of masculine sexual behavior as appearing in male animals, and feminine activity in females; but there are numerous observations to prove that normal, untreated female

mammals are capable of exhibiting nearly all elements of the masculine copulatory pattern, and that sexually active males occasionally display feminine coital responses. Animals showing mating activities typical of the opposite sex also execute the copulatory pattern of their own sex with normal frequency and effectiveness, and these facts must be taken into consideration in any analysis or interpretation of physiological mechanisms responsible for courtship and coition.

a. *Incidence and characteristics.* The display of male-like mounting with pelvic thrusts by normal females has been described in the case of the rat (29, 36, 51), guinea pig (331), rabbit (158), golden hamster (reviewer's observation), porcupine (265a), short-tailed shrew (232), marten (204), dog and cat (reviewer's observation), sow (3), cow (327), rhesus monkey (76, 149), and chimpanzee (55, 325). Records of the appearance of feminine sexual behavior in male mammals are less common, but estrous reactions have been observed in male rats (29, 47, 49, 289), guinea pigs (60), golden hamsters (reviewer's observation), rhesus monkeys (75, 149), and chimpanzees (55).

It is noteworthy that male rats which exhibit feminine sexual reactions invariably prove to be unusually vigorous copulators when placed with receptive females (29, 47, 289). They initiate coition with very little delay, execute few incomplete copulations and reach the point of ejaculation much sooner than most animals of the same strain. Furthermore, such rats are most likely to display female behavior at times when they are highly aroused as a result of previous contact with estrous females. It has been my experience that the most reliable method of eliciting the feminine mating pattern is to allow the male several copulations with a receptive female, to remove the female before the male can ejaculate, and to introduce a second male who will mount and palpate the bisexual individual. If the experimental animal is not sexually aroused, or if he has just ejaculated, it is difficult if not impossible to evoke feminine responses.

The ear-vibration, which is an integral part of the female rat's sexual pattern, may be executed during copulation by male rats that are highly aroused but incapable of ejaculation (50). The males which have been observed to show this reaction were individuals castrated at birth and injected with large amounts of androgen in adulthood. As a result of very early castration the penes were very small and did not permit intromission at normal frequency; but androgen treatment resulted in a high level of sexual excitability, and the experimental animals executed numerous copulatory attempts when placed with females in heat. Although the males gave many signs of being intensely excited, the infrequency of intromission prevented ejaculation; and under such conditions some individuals repeatedly displayed ear vibrations while they were mounting and palpating the female.

The foregoing observations suggest that the neuro-muscular mechanisms mediating feminine sexual reactions in the male are most likely to be activated when a state of pronounced sexual excitement exists, and when execution of the complete masculine pattern is prevented. In addition, of course, application of the appropriate pattern of external stimulation is essential if the entire sequence of feminine responses is to be elicited.

In connection with the question of external stimulation it should be observed that the masculine mating pattern is called forth in females by the same stimuli which evoke it in males, while the feminine responses of males are elicited by those external events which normally initiate these reactions in the female. Thus, female rats exhibit masculine mating behavior most frequently and intensely when they are confronted with other females which are highly receptive (51); and males of this species show feminine reactions when they are mounted and palpated by other males (47).

In considering these facts one becomes aware of the possiblity of dealing with courtship and copulatory behavior in terms of stimulus-response relationships which are relatively constant in all individuals regardless of genetic sex; and it is apparent that the neuro-muscular mechanisms responsible for many such relationships are present in both male and female mammals. In either sex the application of pressure to the dorso-lumbar region tends to elicit lordosis and opisthotonus; and in either sex the multi-sensory pattern of stimulation provided by the receptive female tends to evoke pursuit, mounting and palpation with pelvic thrusts.[17]

b. *Limitations*. It would be a mistake to minimize the limitations of potential bisexuality in mating behavior. Briefly stated these are three in number: (1) limitations imposed by differences in the ovarian and testicular hormones, (2) limitations imposed by genetic differences in the thresholds of responsiveness within the neural mechanisms for masculine and feminine behavior in the two sexes, and (3) limitations imposed by differences in morphological structure, particularly that of the genitalia.

As will be explained at a later point in this review, sensitivity to external stimulation within the neuro-motor circuits mediating masculine sexual responses is greatly increased by testicular hormone, while mechanisms for feminine responses are rendered sensitive by ovarian secretions. Therefore the masculine mechanisms in the genetic female are relatively insensitive because of the absence of the appropriate hormonal facilitation, and a similar situation obtains for the feminine mechanisms in the male.

Even in the absence of any gonadal hormones there is a difference in the ease with which masculine and feminine reactions can be elicited in male and female mammals. For example, the male rat castrated on the day of birth (50), and the male guinea pig subjected to similar operation long before puberty (275) are known to display rudimentary masculine sexual reactions when placed with the receptive females during adulthood; but such animals do not exhibit feminine copulatory behavior when mounted by other males. Thus it appears that in the genetic male the threshold of responsiveness in the neural circuits for mascu-

[17] The suggestion that the genetically-organized neuro-muscular system is, within limits, potentially bisexual, carries no implication of antagonism between the feminine and masculine mechanisms, and should not be confused with such nonsensical formulations as the following: "Thus one must assume that the cerebrum contains male and female centers whose antagonistic action and relative strength determine the individual's sex behavior. Homosexuality results from the victory of the wrong center" (244).

line coital performance is inherently lower than the threshold in the mechanisms for feminine behavior. Whether or not the inverse situation obtains in the female is not known, but a quantitatively similar differential may be assumed to exist, although it seems likely that the masculine circuits in the genetic female are more easily stimulated than are the feminine mechanisms in the male (51).

When a spayed female rodent is injected with androgen, and the neuro-muscular mechanisms associated with masculine sexual behavior are rendered highly sensitive to exteroceptive stimulation, the resulting behavior is less complete than that of the normal male. The difference appears to derive at least in part from the absence of masculine genitalia. Although she mounts other females in the same manner as a male, and executes numerous copulatory thrusts, the androgenized female executes relatively few complete copulations and does not show the ejaculatory pattern (15). In fact her behavior resembles closely that of adult males from which a portion of the penile bone has been removed (50). These observations make it possible to assume that the incompleteness of the pattern is in both instances referable to lack of intromission with consequent failure of one important source of sexual stimulation which operates to produce intense excitement in the normal male.

5. *Behavior involving auto-genital stimulation.* a. *Incidence and characteristics.* Under certain conditions male and female mammals of many species manipulate their own external genitalia. In some animals auto-genital stimulation occurs in close conjunction with copulatory behavior, under the influence of sexual excitement generated by previous contact with the opposite sex. Members of other species exhibit similar responses in isolation, and their behavior indicates that the resulting sensations induce definite erotic effects.

Although there have been no experimental attempts to determine the physiological mechanisms involved, the widespread occurrence of auto-genital manipulative reactions suggests that this type of behavior may be a fundamental element in the basic mammalian sexual pattern, and as such it deserves attention in the present review.

It is reported that sexually excited male porcupines manipulate the genitalia with the forepaws and often walk about on three legs, straddling a long stick which is held in one front paw and extends backward between the hind legs, coming into intimate contact with the penis (265a).

Licking and manipulation of the penis constitute integral parts of the coital pattern in many rodents, including the rat (286), golden hamster (247b), and short-tailed shrew (232). In the last-named species post-copulatory penis manipulation is thought to be essential to fertile mating, for after each mount regardless of the occurrence of intromission the male must return the penis to its sheath before he is capable of pursuing the female (232).

Ejaculation has not been observed in connection with auto-genital manipulation by male rodents and it is not at all certain that such activity is sexually stimulating, but there is some evidence to suggest this may be the case. According to some experimenters the male rat's penis licking often is accompanied by distinct pelvic movements similar to those occurring during copulation (286);

and males of this species in which overt copulation has been inhibited by electric shock respond to the presence of the estrous female with a pronounced increase in the frequency and duration of oral manipulation of the phallus (314).

Females of lower mammalian species characteristically nibble and lick the vaginal orifice after each copulation or series of copulations, and this response forms a particularly prominent element in the post-copulatory "after reactions" shown by female cats. The female porcupine in heat may seize and straddle small sticks, supporting one end with a forepaw and riding the stick about the cage in such a manner as to stimulate the external genitalia (265b).

Caged male rhesus (177) and cebus (129) monkeys often masturbate to the point of orgasm, and such behavior has sometimes been interpreted as an "abnormal" response induced by artificial conditions of sexual deprivation (149). This point of view appears to reflect moralistic conviction rather than accurate observation, for masturbation with ejaculation occurs in free-ranging, socially dominant male macaques which have free access to receptive females (76). Adult male chimpanzees frequently stimulate ejaculation by manual or oral masturbation (276), and immature males of this species masturbate several years before ejaculation is possible (55).

Auto-genital stimulation is not peculiar to male primates, for it occurs in the female rhesus (177) and in both infantile and adult chimpanzees (55) and gorillas (323). The female ape may masturbate with the fingers or with foreign objects such as fruit, pebbles, leaves, or projecting corners of boxes.

In humans as in other primates auto-genital manipulation begins long before puberty and occurs in a high proportion of the population. According to one authority, "masturbation with partial orgasm may begin in the first year of life and continue irregularly thereafter in both sexes" (208). Other observers state that grasping and manipulation of the penis are frequent by the third year in the human male. although genital stimulation is said to occur less often in the female (190).

Regardless of the exact age of its initial appearance (which must be subject to considerable individual difference) there is no doubt concerning the very high frequency of masturbatory behavior in reproductively immature humans, particularly males. In one series of interviews 90 per cent of 291 boys reported the practice of masturbation beginning before or soon after the onset of adolescence (246). That activity of this character is not limited to members of European cultures is amply demonstrated by reports of its widespread occurrence among primitive tribes (122).

b. *Establishment of masturbatory habits in the human.* In view of the common tendency to regard human masturbation as an "abnormal" or "substitute" response it may be worth while to point out one course of events by which deep-rooted habits of genital manipulation may be established during infancy or early childhood.

It is well known that in most if not all primates genital reflexes appear soon after birth. In the macaque monkey penile erection is seen during the second post-natal week (166), and penile and clitoral tumescence occur in chimpanzees

at birth (55). Male and female human infants exhibit turgescence of the penis or clitoris at birth, and in the case of males three to twenty weeks of age the average frequency of erection varies from four to thirty-five within a 24-hour period (148). The occurrence of a definite response to masturbation in human infants is indicated by the report that manipulation of the erect penis or clitoris is followed by "unmistakable relief from tension or restlessness" (208). Although the limited neuro-muscular capacities of very young infants would not permit prolonged auto-genital stimulation, it is known that grasping and handling of the phallic organs occur frequently by the third year (190).

To these data concerning genital activity must be added the fact that conditioned reflexes can be established in very young infants. For example, if a buzzer is always sounded just before presentation of the breast or nursing bottle, babies less than one week old react to the auditory stimulus by opening the mouth and making sucking movements (205).

It may be suggested that the spontaneous occurrence of genital tumescence tends to elicit manipulatory responses, that these in turn result in detumescence accompanied by relief of tension, and that the resultant stimulus-response association is strengthened by each repetition. In such a case regular habits of auto-genital stimulation could become firmly established in infants as soon as neuro-muscular development was sufficient to permit localization and repetitive manipulation of the genitalia.

HORMONAL FACTORS IN SEXUAL BEHAVIOR. Many observational and experimental findings combine to indicate that various hormones, particularly those of gonadal origin, may exert pronounced effects upon the sexual behavior of male and female mammals. Earlier reviews provide comprehensive summaries of the literature on this topic (49, 86, 91, 288), but the pertinent data will be surveyed in a somewhat different manner in the present article. In preceding sections an attempt has been made to define the neuro-physiological factors responsible for courtship and mating, and it has been stated in passing that the functional capacity of certain nervous mechanisms is influenced by endocrine products. It now becomes necessary to consider the evidence upon which such conclusions are based and to attempt to understand the ways in which glandular secretions may exert their specific behavioral effects.

A. *Relation of Gonadal Cycles to Mating Activity.* 1. *Normal cycles in lower mammals.* In males and females of seasonally breeding species and in females of polyestrous forms complete mating performance normally occurs when the secretory activity of the gonads is at or near the maximum (49). The heat responses of female rats (59) and guinea pigs (329) usually appear only when mature ovarian follicles are present, and a similar correlation exists in the case of other domesticated animals such as the cow (154) and horse (153). Males of these species are continuously active sexually and the secretory activity of the testes appears to be more or less constant; but in seasonally breeding forms like the marten (204) and deer (319) the male's rutting behavior coincides with the recrudescence of the testes.

Although a close positive relationship between high gonadal activity and overt

sexual behavior is a general rule in nearly all lower mammalian forms, some individual exceptions exist in most if not all species. Some female mink whose ovaries contain many large follicles refuse to receive the male (109);[18] and a few rats (59), guinea pigs (61) and cows (155) show normal ovarian cycles but never exhibit heat behavior. All strains of laboratory rats apparently contain an occasional male that never displays copulatory activity despite the possession of histologically normal testes and well developed accessories (40).

In the majority of sub-primate species the pregnant female is not sexually receptive, but exceptions to this rule are evidenced by reports of frequent copulation shown by gravid golden hamsters (62) and rabbits (158). In the latter species, as a matter of fact, willingness to receive the male bears no observable relation to fertility, for it has been found that in the wild rabbit, *Oryctolagus cuniculus*, males copulate twice as often with pregnant or diestrous females as with those which are capable of insemination (63a).

2. *Induced cycles in lower mammals.* By the administration of gonadotrophic substances it is often possible to induce unseasonal development of the gonads, and the consequent increase in gonadal hormones results in the appearance of mating responses in female ferrets (130), sheep (85), swine (332), horses (90a, 93a, 157), cattle (154), cats (126, 219), dogs (189a) and rodents (112, 320). It should be added, however, that although follicle growth and ovulation occur in a high percentage of females subjected to gonadotrophic stimulation, the induction of estrous behavior is achieved in a smaller proportion of cases (235a).

Gonadotrophic stimulation of the immature testis or ovary may evoke the secretion of endogenous gonadal hormones which in turn cause precocious sexual activity. For example, female mice injected with fresh pituitary substance will mate at fifteen days of age, and rats subjected to the same treatment show all of the behavioral signs of estrus at twenty-two days although at this age they are so small that adult males cannot effect copulation (272). Female rats injected with chorionic gonadotrophin on the twenty-second day after birth mate and conceive three days after treatment (83), and administration of anterior pituitary extracts to immature male rats is reported to result in "premature psychic maturation" (283).

Findings of this nature indicate that the neuro-muscular factors involved in adult sexual behavior are functionally organized long before the advent of reproductive maturity, and that the activation of these mechanisms merely necessitates the presence of adult hormone levels. As a matter of fact there is one bit of evidence to show that in some species adult mating responses are possible immediately after birth. For the first few post-natal hours female and male guinea pigs respond to tactile stimulation of the rump with the execution of lordosis and opisthotonus in a manner typical of the sexually-receptive adult female (60). Presumably these reactions occur under the influence of temporarily surviving maternal hormones; in any event they disappear within the first day after birth.

[18] In this species there appears to be little direct correlation between waves of developing follicles and willingness to copulate (109).

3. *Normal cycles in primates.* The interrelationship between ovarian cycles and sexual behavior is less clear-cut in the intermediate and higher primates than in lower mammals. In more primitive primates such as the lemuroids the female's receptivity apparently is still heavily dependent upon gonadal hormones, for the female galago, *G. senegalensis mohli* will mate only during physiological estrus (193a). Among the monkeys, however, the situation is altered slightly, for although the female rhesus is fully receptive only at the time of ovulation (20) she may permit copulation during other stages of the cycle (75), and her readiness to accept the male is affected by social and behavioral factors as well as by her own physiological status (76). Even greater independence from ovarian conditions is to be seen in the sexual activities of female apes. The chimpanzee may display sex presentation and receive the male at any point in the menstrual cycle (309, 324, 326), and although marked individual differences undoubtedly exist, it appears that this primate stands midway betwen the monkey and human as regards the degree to which ovarian condition affects social behavior (330).

It has proven extremely difficult to identify possible effects of ovarian cycles upon sexual responsiveness in the human female, but it seems fairly certain that hormonal influences are much less important in our own species than in any lower form. Various investigators have reported that most women experience maximal sexual excitability just before and just after the period of menstrual flow (93, 96, 299), but non-hormonal factors undoubtedly contribute materially to any such variation in desire for intercourse.

Incomplete as they undoubtedly are, the available data strongly suggest the existence of an inverse relationship between advancing phyletic status and the degree to which the female's sexual behavior is controlled by ovarian hormones. The full significance of this observation becomes apparent when it is considered in conjunction with previous statements concerning the evolutionary increase in the importance of cerebral functions to the sexual activities of mammals.

B. *Effects of Gonadectomy.* Direct evidence for the function of testicular and ovarian hormones in sexual behavior is provided by descriptions of behavioral changes which follow removal of the reproductive glands.

1. *Removal of the ovaries.* The effects of ovariectomy upon mating behavior vary according to the phylogenetic status of the species, and in higher mammals individual differences are apparent. In female rodents (165, 296, 315), ungulates (7, 211), and carnivores (22, 181, 189, 197) the operation occasions prompt and complete elimination of copulatory reactions.

In the female rhesus monkey sexual responsiveness ordinarily is lost rapidly after ovariectomy (13), but under special conditions, such as those consequent to specific types of brain injury, copulation may occur frequently despite gonadectomy (177). Female chimpanzees tend to show sexual presentation and permit copulation many months after loss of the gonads, although the residual behavior is less frequent and prolonged than that occurring during the ovulatory period in intact animals (80).

The sequelae to human ovariectomy are in dispute. According to different authors this operation may increase (203), decrease (81, 124, 162, 202), or have no effect upon sexual desire in women (115, 312). The lack of agreement between

authorities strongly suggests that individual differences are marked and that hormonal effects are much less powerful and less direct in the sex life of the human female than in that of lower forms.

2. *Removal of the testes* The effects of castration upon sexual behavior in male mammals differ with the species, with the individual, and with the age of the animal at the time of operation.

a. *Pre-puberal castration.* It is sometimes stated that male animals gonadectomized before puberty exhibit no mating behavior (86), but the limited applicability of this generalization is indicated by reports of frequent if incomplete copulatory reactions in pre-puberally castrated male rats (38) and guinea pigs (218, 265, 275). The responses observed in such animals usually consist of mounting the receptive female and exhibiting pelvic thrusts. Erection does not seem to occur very often, intromission is infrequent, and ejaculation is lacking (50). Pre-puberally castrated male chimpanzees show much more sexual activity than do similarly operated rodents, and at least some adult apes deprived of testes in infancy are capable of frequent copulation with erection and intromission (79).

b. *Castration in adulthood.* Male mammals from which the testes are removed in adult life may continue to display coital reactions for many months after operation. Casual observation has shown that copulatory behavior survives castration in male horses (277), cattle (307), and monkeys (308), although no data have been reported which would indicate the completeness or intensity of the residual sexual activity. Quantitative studies of pre- and post-castrational mating performance in male rats (46, 268, 277, 291, 296) and rabbits (293) reveal that ejaculation is the first element of the pattern to disappear following gonadectomy, that complete copulation with intromission undergoes a progressive decline in frequency, and that in some males of both species incomplete copulatory attempts persist for several months.

The changes which follow castration in adulthood appear to include a reduction in the drive to seek out the sexually-receptive female, for gonadectomized male rats are less willing than intact animals to cross an electrified grid interposed between them and the estrous female (226); and castrated male guinea pigs display a weakening of the tendency to cross a mechanical hurdle with a female in heat on the other side (265).

The plethoric and contradictory literature dealing with sexual ability in castrated men permits only two generalizations. The range of individual variability is definitely wider in humans than in any other mammalian species (182), and the importance of testicular hormone for susceptibility to arousal and capacity for complete response is less marked in man than in lower forms (216). Some authorities adhere to the extreme view that testis secretions are entirely dispensable, and that "pre-pituitary" and "psychic centers" control libido and potency in the human male (113). The majority of writers, however, concur in the belief that, although their effects may be less important in human sex life than in that of lower animals, gonadal secretions do play a major rôle in the normal sex behavior of man (240, 285, 302).

References in the medical literature report complete survival of sexual powers

in various individual cases studied nine, twenty (257), and thirty (113) years after loss of the testes. In other instances gradual loss of desire and capacity for intercourse is said to have become evident at different intervals from two months to seven years after operation (87, 210, 213). Some castrates report that although the ability to achieve orgasm is reduced or eliminated, sexual desire and the capacity for erection are not affected; and it will be recalled that in the case of male rats and rabbits the ejaculatory response was found to be the most sensitive indicator of post-operative hormonal deficiency (50, 293, 296).

Partial or complete retention of sexual powers after castration might be interpreted as a result of androgenic stimulation provided by hormones of the adrenal cortex; but it is doubtful that any such explanation is universally applicable, for male castrates with very low titers of urinary androgen may perform coitus several times weekly as long as thirteen years after loss of the testes (152).

A major obstacle to estimation of the effects of castration upon sexual activity in men arises from the widely recognized influence of non-hormonal factors which may affect susceptibility to arousal and capacity for intercourse. Repeated references to the power of "suggestion" and "psychological factors" reflect a general acceptance of the fact that in the human as in no other animal the cortical component is of primary importance in the control of sexual phenomena.

C. *Effects of Administering Gonadal Hormones.* Observed correlations between cycles of secretory activity within the gonads and periodic appearance of sexual behavior, and the recognized effects of castration upon mating tendencies lend considerable support to the belief that ovarian and testicular hormones are in some fashion responsible for courtship and mating in many species. Even more direct evidence for such a conclusion is to be seen in the behavioral changes which follow the administration of gonadal hormones to male and female mammals.

1. *Ovarian hormones in the female.* Many reports describe the appearance of sexual receptivity in spayed or anestrous female mammals consequent to the administration of ovarian hormones. Injections of estrogen evoke mating responses in the anestrous cow (157, 321) and ewe (84, 155, 156), and in the spayed bitch (181, 189), cat (22, 197), ferret (206), ewe (21), cow (1a), and rhesus monkey (13). Administration of estrogen to spayed chimpanzees causes swelling of the sex skin and intensifies the tendency to court and cohabit with the male (80). It may be of considerable significance that the intensity and completeness of the sexual pattern induced by exogenous estrogen is related to the amount of hormone administered. For example, spayed heifers receiving estrogen exhibit parts of the pattern after one or two treatments, but continued dosage is essential if all of the normal pattern is to be evoked (1a).

Ovariectomized females of some species, including the guinea pig (58, 61, 165), golden hamster (123), mouse (252), and rat (39), show sluggish heat reactions when treated with estrogen, and normal estrous behavior appears only if estrogen administration is followed after a suitable interval by the injection of progesterone. In other species progesterone appears to act as a sex depressant. Thus, female rabbits in estrus may refuse to mate following progesterone admin-

istration (200), and the estrogen-induced heat of the spayed rhesus female is reduced by injection of progesterone (19).

Individual differences in the behavioral response to hormone treatment have been noted in several species. For example, spayed ferrets vary with respect to the amount of estrogen which must be injected to obtain full behavioral estrus (206), and the same is true in the case of cattle. In general it appears that individuals which exhibit the most intense and complete sexual activity prior to ovariectomy respond post-operatively to lower concentrations of exogenous hormone than do other females whose pre-operative mating reactions were less easily elicited. In one experiment the mating performance of a group of heifers was observed before castration, and when estrogen was given post-operatively it was found that, "in general . . . the easier it had been to detect heat [pre-operatively] the lower was the level of hormone necessary to induce heat after the operation" (7).

It has been noted above that there are in various species a few individual females which show normal ovarian cycles but never come into behavioral estrus. When female guinea pigs of this type are ovariectomized and injected with ovarian hormones sexual behavior may be elicited if very large doses are employed (61), and some heifers that have never shown heat behavior will do so following the administration of very large amounts of estrogen (157). Findings of this nature make it possible to assume that the original failure of normal sexual receptivity was due, not to a sub-normal supply of gonadal hormone, nor to incompletely organized nervous mechanisms, but to the fact that the neural circuits lacked normal sensitivity to the facilitating effects of physiological concentrations of the essential endocrine secretions.

There are certain observations which indicate that comparable individual differences are to be found among female chimpanzees. In one study (80) four ovariectomized apes were treated with estrogen and three animals became very active sexually, showing mating behavior identical to that of intact females at the height of genital swelling. The fourth female exhibited the expected swelling of the sexual skin but would not receive the male; and consultation of breeding records revealed that prior to ovariectomy this particular individual had never permitted males to copulate.

The fact that in the human female sexual responsiveness is correlated loosely if at all with cycles of estrogen production, and the prevalence of contradictory reports concerning the sequelae to ovariectomy in women would not lead us to expect any general agreement as to the effects of hormone treatment in ovariectomized women. However, it is reported that following this operation some women continue to display satisfactory marital adjustments only if estrogen is administered (135). In one study seven oophorectomized women received estrogen. Five patients experienced a restitution of libido which had been lost post-operatively, one case in which pre-operative responsiveness had been low described a marked gain under treatment, and the seventh patient reported post-operative loss with no improvement under hormone (161). Attempts to increase post-menopausal responsiveness by hormone therapy have not been

successful to any degree, and in general clinical results suggest that any concept of human sex drive as solely dependent upon ovarian secretions is "an over-simplification of an extremely complicated and subtle reaction" (267). The major importance of psychological factors is strongly indicated by reports of sexual rehabilitation resulting from psychiatric treatment (91).

In connection with the depressing effects of progesterone upon estrous behavior in some lower mammalian species it is of interest to note that this hormone is said to decrease the desire for intercourse in women (146).

2. *Testicular hormone in the male.* Androgen administration results in the restoration of normal copulatory behavior in pre- or post-puberally castrated male rabbits (125), guinea pigs (219, 265, 275), and rats (219, 266). The progressive change in sexual behavior which accompanies daily treatment of castrated male rats is precisely the reverse of the modifications that follow removal of the testes. Initial injections evoke an increase in the frequency of incomplete matings. Continuation of hormone treatment calls forth complete copulation with intromission, and lastly the ejaculatory response is restored (46, 50, 296). Data of this sort illustrate the important point that the intensity and completeness of sexual behavior depend in part upon the level of gonadal hormones, a conclusion corroborated by reports of several investigators (71). For example, although in the absence of hormone treatment the pre-puberally castrated male chimpanzee may display frequent copulation with intromission, the administration of androgen is necessary to the occurrence of ejaculation (79).

It is abundantly clear that the concentration of testis hormone just adequate to elicit a given mating response in castrated males varies from one animal to another, and individual differences of this sort are definitely related to sexual responsiveness prior to operation (50, 296). In tests conducted before hormone treatment pre-puberally castrated male rats exhibit wide differences in responsiveness to the receptive female; and when testosterone propionate is administered individuals that were highly reactive prior to treatment tend to copulate at lower dose levels than do those which showed little or no sexual activity in the absence of androgen (46). This observation suggests that males which respond to comparatively small amounts of hormone are individuals in whom the nervous mechanisms for sexual arousal are relatively excitable even when hormonal facilitation is lacking.

Among males as among females there are some individuals who fail to show sexual behavior although the histological condition of the accessory organs indicates that gonadal hormones are secreted in normal amounts. Male rats of this type sometimes copulate when the concentration of endogenous hormone is augmented by administration of androgen (31, 294), and this suggests that the original failure to mate may have been due to relative insensitivity to normal hormone level on the part of the neural mechanisms involved in arousal and manifestation of sexual excitement.

It has been recorded that some human males suffer noticeable reduction in sexual capacity and desire following loss of the testes; and in many such cases the administration of androgenic hormone results in increased frequency of erection,

intensification of sexual drive, heightened capacity for intercourse, and the appearance of erotic dreams (73, 141, 303). It is worthy of note that in men as in the males of other mammalian species the degree to which normal sexual function is restored varies with the dose level of the hormone (150, 241).

3. *Response of immature animals to gonadal hormones.* At an earlier point it was noted that pre-pubescent males and females may display adult copulatory responses when the immature gonads are stimulated to precocious secretory activity by the administration of gonadotrophins; and here it can be added that similar behavioral changes are evoked by direct application of gonadal steroids. Sexual mounting appears in male rats as early as sixteen days after birth when androgen is administered on the fourteenth day; and females injected with estrogen show lordosis at twenty-one days, although in the normal course of development this response does not appear until a month or so later (43).

In like fashion the administration of estrogen to the immature bitch is followed by the appearance of all signs of estrus including willingness to accept the male (189). Copulatory behavior has not been observed, but there is proof that the sexually-stimulating qualities of the very young female chimpanzee are markedly increased by estrogen treatment. Although adult males show little sexual interest in normal infantile females, proximity to an estrogen-injected female less than two years old is reported to have elicited intense excitement in an adult male as evidenced by the prompt occurrence of erection followed by masturbatory responses with ejaculation (104).

The facility with which testicular and ovarian hormones stimulate adult copulatory performance in immature rodents points again to the conclusion that the neuro-muscular organization for such reactions is completed long before mating behavior normally appears, and that the absence of any feminine responses in pre-adolescent females and the incomplete nature of the pre-pubescent male's sexual activity are due primarily to lack of appropriate hormonal facilitation.

Although the mechanisms for mating appear to be completely organized relatively early in life, and although they frequently can be rendered functional under the influence of exogenous hormones, their sensitivity to hormonal effects is not fully developed during the period of infancy. On the contrary, behavioral responsiveness to endocrine secretions appears to be low in the very young infant and to increase post-natally in much the same fashion as does the sensitivity of such other elements in the reproductive complex as the uterus, vagina, prostate and seminal vesicles (243).

New-born guinea pigs respond to tactile stimulation of the rump with the exhibition of lordosis and opisthotonus in a manner typical of the adult estrous female (60), and this indicates that the essential neuro-muscular units are fully organized pre-partum. Presumably the post-partum heat responses, which persist for only a few hours, are due to the temporary effects of maternal hormones. In any event it has been shown that females seven days old fail to display heat responses when injected with ovarian hormones, even though the dosage be very high. Examinations repeated at regular intervals show a rapid increase in

sensitivity to the hormones with maximal responsiveness appearing at about thirty days after birth and continuing for two years (316). The progressive change in reactivity is not due to endogenous ovarian secretions, for it occurs in individuals spayed at birth. The female rat is similarly refractory to ovarian hormones during infancy, and in this species maximal sensitivity is also attained about the thirtieth day after birth (316).

I have suggested quite tentatively that the absence of mating behavior in pre-puberal females of lower mammalian species, and the relative infrequency and incompleteness of copulatory responses in males of the same species are due, not to nervous system deficiencies, but to lack of appropriate hormonal facilitation. Now if, for the purpose of exposition, this view is accepted, it provides supplementary evidence in connection with previously-mentioned differences in the extent of pre-puberal sexual behavior shown by primates and non-primates. There is, so far as I am aware, no reason to believe that the testes and ovaries reach an adult level of secretory function at an earlier developmental stage in primates than in non-primates; and accordingly the more complete copulatory activity of immature male and female monkeys, apes and humans may be interpreted as a function of the relatively greater importance of cerebral factors in the sexual life of higher forms.

D. *Hormonal Factors in Bisexual Mating Behavior.* In a preceding section of this review it was pointed out that under certain conditions male and female mammals may display portions of the mating pattern of the opposite sex. Such behavior has commonly been "explained" as a product of heterologous gonadal hormones, and rarely has sufficient emphasis been placed upon other variables which are equally powerful in determining the type of sexual activity that will occur in a given situation. From what has been said previously it should be apparent that the appearance of a particular mating response is affected by the pattern of external stimulation and by the functional characteristics of the neuro-muscular constitution as well as by the chemical nature of the facilitating hormonal agents. Only when all of these factors are taken into account is it possible to formulate an adequate interpretation of the execution of masculine reactions by genetic females and feminine responses by males.

At this point it will be profitable to consider first the hormonal interpretations of bisexual behavior, and then to consider a more comprehensive analysis based upon less restricted views of the physiological basis for such phenomena.

1. *Concepts of specificity and antagonism in the behavioral effects of ovarian and testicular hormones.* Evidence derived from a variety of sources has led to the formulation of theories which hold that the effects of gonadal hormones are highly specific, that only ovarian hormones can stimulate female morphological and behavioral characters, while testicular secretions are solely responsible for masculine bodily structures and overt reactions (192, 217).

a. *Spontaneous or experimental gonad reversal.* Transformation of an ovary into a testis and the converse change are fairly common occurrences in members of some sub-mammalian classes (82, 89), and morphological alterations of this type are sometimes accompanied by corresponding reversal of the overt sexual behavior. Experiments with mammalian material suggest that surgical ex-

change of masculine and feminine gonads may induce like results. Ovariectomized does bearing implanted testes behave sexually like the mature buck (191), and castrated male rats (278, 279, 280) or guinea pigs (274, 282) into which ovaries have been grafted exhibit estrous responses typical of the females of their species. Spayed female guinea pigs containing active testicular grafts respond to normal estrous females by displaying many elements of the masculine mating pattern (215, 282).

b. *Administration of heterologous gonadal hormones.* Changes which are reported to occur in mating behavior following the administration of androgen to females and estrogen to males lend support to the belief that gonadal secretions exert highly specific influences upon sexual activities.

The administration of androgenic preparations occasions the appearance of certain masculine copulatory responses in female rats (15, 180, 295), rabbits (170), dogs (54), and sheep (84); and castrated male rats treated with estrogen show various elements of the female mating pattern (14).

This sort of evidence indicating a high degree of specificity in the behavioral effects of ovarian and testicular hormones is buttressed by additional data which suggest that the execution of masculine sexual reactions is inhibited by estrogen, while androgen tends to depress feminine mating responses (282). It is stated, for example, that the responsiveness of male dogs to the bitch in heat is lowered by estrogen treatment (222); and men given estrogenic substances are said to experience a decrease in sexual desire (100, 121, 220).

2. *Evidence against concepts of antagonism or complete specificity.* Although there is no doubt that the behavioral effects of androgen and of estrogen are quite distinguishable and relatively specific there are definite limits to such specificity; and theories of antagonistic action are definitely questionable.

a. *Mating behavior of intersex individuals.* Intersex pigs with feminine external genitalia and lateral hermaphroditismus always attempt to copulate in masculine fashion although they may display some of the signs of estrus (26); and hemi-castrated male guinea pigs with one testis and one implanted ovary mate in masculine or feminine fashion, "according to the animals with whom they are brought into contact" (259, 281). I have recently discovered in my animal colony several Norway rats in which the external genitalia consist of a greatly hypertrophied clitoris and a penetrable vaginal opening. The internal reproductive organs of some of these individuals include one ovary and an ovitestis. Mating responses are bisexual, and either male or female copulatory patterns may appear under appropriate external stimulation.

The sexual behavior of intergrade pigs and of naturally-occurring or experimentally-produced hermaphroditic rodents shows that individuals bearing testicular and ovarian tissue may display overt mating reactions of both sexes, and such observations are difficult to reconcile with any concept of antagonism between the behavioral effects of ovarian and testicular hormones. How then are we to interpret the reported inhibition of masculine sexual behavior by estrogen administration, and the diminution of feminine mating reactions by androgen?

In the first place, none of the experimental reports describing results of this

nature include satisfactorily objective or quantitative behavioral records to support the conclusions advanced; but even if such data were available, the claim for an antagonistic effect of heterologous hormones would not be justified. A more reasonable interpretation rests upon the established fact that administration of exogenous gonadal steroids tends to depress the gonotrophic functions of the anterior pituitary and thus to interfere with the production of endogenous testicular or ovarian secretions (217).

The explanation proposed here would account for an observed decrease in sexual behavior following treatment with heterologous hormone as a result of lowered concentrations of endogenous gonadal agents. Support for this interpretation is found in reports that the implantation of ovarian tissue in the intact male guinea pig is followed by a decrease in the secretion of testicular androgen (270). Furthermore, men in whom sex drive has been lowered by estrogen treatment may experience a revival of libido when gonadotrophic substances are administered (142),—a result strongly suggestive of initial interference with hypophyseal stimulation of the testes.

b. *Bisexual reactions in morphologically normal animals.* Even more convincing evidence against theories of extreme specificity or antagonism in the behavioral effects of gonadal secretions is to be found in descriptions of masculine mating responses shown by morphologically normal female mammals and, less frequently, the display of feminine sexual reactions by untreated males.

At an earlier point in this review I have referred to the male-like mounting and execution of pelvic thrusts shown by female rats (18, 29, 36, 51, 163, 180), hamsters (reviewer's observations), guinea pigs (329, 331), rabbits (158), porcupines (265a), short-tailed shrews (232), marten (204), dogs and cats (reviewer's observations), African lions (88), swine (3), cattle (327), rhesus monkeys (76, 149) and chimpanzees (55, 325). Some writers have suggested that this type of behavior in females is caused by androgen (163), possibly of ovarian origin (284). However such suggestions seem to be based upon no more than an *a priori* assumption that the appearance of masculine sexual responses demands the presence of androgen.

There are several lines of evidence which seem to argue against the theory that masculine coital responses in the female are a product of ovarian androgen. For example, spayed heifers display mounting behavior (which is a common element in the normal estrous pattern of this species) when they are injected with estrogen (1a). The male-like mounting responses of the female guinea pig are dependent at least in part upon ovarian secretions, for the frequency of such reactions is correlated with the number of maturing follicles (329). In addition, this behavior is induced in the spayed female by the administration of estrogen and progesterone; the same hormones which elicit sexually-receptive reactions (331). However, although the female's masculine sexual activity and feminine receptivity seem to be controlled by the same endocrine substances, it is important to note that in some individuals the male behavior appears in the absence of heat (328); a finding which has been taken to indicate that different nervous mechanisms are involved (327). Ovarian secretions are not essential to mas-

culine copulatory responses in the female rat. There is no relationship between such behavior and willingness to receive the male, and the frequency of male reactions is not affected by pre- or post-puberal ovariectomy (51).

Feminine mating reactions are sometimes shown by male rats (29, 289), hamsters (reviewer's observations), rhesus monkeys (149) and chimpanzees (55). Only in the rat has the hormonal basis for such behavior been studied (47). The single animal described responded to receptive females with prompt and vigorous copulation, and reached the point of ejaculation much more quickly than the majority of males of the same strain. When mounted and palpated by other males the experimental animal shifted immediately to the feminine coital pattern. The feminine reactions were intense and complete and included hopping, ear-vibration and deep lordosis.

When the male was castrated the feminine mating reactions disappeared in less than a week, while masculine sexual reactions underwent the gradual decline which is characteristic in the adult castrate. Subsequent administration of androgen revived masculine copulation and ejaculation, and restored most of the feminine reactions although the latter were sluggishly executed. Injections of estrogen and progesterone evoked the complete feminine pattern as well as most of the masculine responses, but the ejaculatory reflex was lacking.

It appears that prior to operation both feminine and masculine sexual patterns were supported by testicular secretions; and the results of post-operative hormone treatment further indicate that androgen was capable of stimulating both types of behavior. The capacity of ovarian hormone to facilitate female behavior is in accord with evidence presented earlier; and the partial revival of masculine reactions under the influence of these substances receives more detailed consideration in the following paragraphs.

c. *Paradoxical effects of gonadal hormones.* Simplified theories proposing complete behavioral specificity of estrogens and androgens are further thrown into question by experimental data which demonstrate that feminine sexual behavior may be increased by administration of androgen and that masculine reactions are sometimes intensified under the influence of estrogenic treatment.

Pre- or post-puberally castrated male rats characteristically show incomplete mating responses, and complete copulations occasionally occur. If estrogen is injected the frequency of both incomplete and complete contacts rises noticeably, although the ejaculatory reaction is not revived (14, 37, 38). Estrogen has been observed to evoke male behavior in females as well as males. Anestrous ewes treated with estrogenic hormone tease and ride other females in masculine fashion (156, 157, 211), and the implantation of estrogen pellets in heifers is sometimes accompanied by such bull-like behavior as roaring, pawing the ground with the fore-hoofs, and pursuing and mounting estrous females (157). The masculine copulatory reactions of these estrogen-treated females include pelvic movements so vigorous that they culminate in a high proportion of pelvic fractures (119, 120). Removal of the hormone pellets is followed within three or four days by disappearance of the masculine sexual responses.

Complementary data are found in reports of the stimulation of feminine

copulatory performance by androgen. Untreated, spayed female rats show no estrous reactions. If androgen is injected into such animals they exhibit strong tendencies toward the masculine sexual pattern in response to receptive females, and, in addition, they display lordosis when mounted by other rats (18, 37, 163). Some clinicians report that lack of sexual desire and low responsiveness in intercourse may be treated more effectively in the human female by androgen than by estrogen (145, 147, 258). Androgen may also facilitate feminine sexual performance in the genetic male, for normal male rats injected with massive doses of testosterone propionate continue to execute masculine coital responses to the estrous female and begin to display feminine reactions when mounted by other males (33).

3. *Hormones and human "homosexuality"*. The term "homosexual" has been employed in so many different frames of reference and with so many different connotations that its scientific validity is questionable, and its use is apt to detract seriously from a clear understanding of human sexual behavior. It would be fruitless in this discussion to attempt any original definition, or to describe the variety of contradictory definitions which have been sponsored by earlier writers. For present purposes it is sufficient to observe that there are human adults who appear to be capable of sexual arousal and genital activity only when stimulated by members of their own sex, and who feel either indifference to or active revulsion for heterosexual contacts. On strictly hypothetical grounds some writers have attempted to account for such behavior in terms of an insufficiency of homologous gonadal hormone and/or a hyper-normal concentration of heterologous secretions (133, 134, 140, 167, 194, 238). This theoretical point of view is difficult to reconcile with the consistent failure of attempts to "cure" the condition by hormone treatment (297), and with the additional fact that administration of androgen to male homosexuals tends to increase the frequency of their sexual activity with other men (134).

Recognition of the obvious contribution of experiential factors to the structuring of human sexual activity raises serious doubts as to the directive influence of gonadal hormones in homosexuality (110, 130). The guiding effects of training and experience are reflected in the fact that the preferred sex rôle and the direction of sexual desire in human hermaphrodites tend to be in accordance with the up-bringing of the individual (as a "boy" or "girl"), and are relatively independent of internal or external reproductive morphology (116). One study of 84 "pseudo" and "true" hermaphrodites revealed a high correlation between the form of overt sexual expression and the manner in which the person had been raised; whereas there was a much less marked relationship between sexual behavior and the character of the gonads. The author interpreted his findings to mean that, "while the *power* of the human sex drive may possibly be largely dependent upon physiological factors, such as the quantity and/or quality of sex hormone secretion, the direction of this drive does not seem to be directly dependent upon constitutional elements" (106).

The potent influence of individual experience is not restricted to overt sexual behavior, but affects equally the more generalized expressions of sex such as those measured in psychological tests of "masculinity" and "femininity."

Application of the Terman-Miles "M-F" test to human sex intergrades revealed that in many cases the final score inclined heavily on the side contrary to biological sex (305).

E. *Rôle of Non-Gonadal Hormones in Sexual Behavior*. The conclusion is inescapable that the secretory products of several endocrine glands must exert indirect influences upon the sexual behavior of male and female mammals, but supporting evidence as yet is scanty.

1. *Anterior pituitary*. Pituitary control of sexual activity appears entirely indirect, being limited to the stimulation of secretory activity within the gonads. This influence, however, is of primary importance; and hypophysectomy eliminates mating in the male rat (250, 274) unless the testes are maintained by administration of gonadotrophic substances (273). The possibility that the gonadotrophins might be directly essential to sexual behavior is contra-indicated by the survival of normal mating reactions in hypophysectomized male rats (150, 225) and female cats (198) treated with testicular and ovarian hormones respectively.

2. *Adrenal cortex*. Steroid hormones secreted by the adrenal cortex are known to possess certain androgenic properties (68), and it has been suggested that some cases of survival of sexual behavior in castrated men may be due to cortical androgens (151). Although it is purely speculative a comparable explanation might account for the masculine copulatory behavior shown by ovariectomized female rats (51). The latter point is one which could and should be attacked experimentally. It will be recalled that progesterone is reported to depress sexual desire in women, and at this point it may be added that desoxycorticosterone acetate is said to exert similar effects (146).

3. *Thyroid*. Thyroid secretions appear to affect sexual activity primarily if not exclusively by way of systemic changes in metabolic function. Male rats display no mating behavior after removal of the thyroid, but thyroidectomized females of this species are capable of fertile breeding (118, 251). This may be compared to similar inter-sexual differences in the importance of the cerebral cortex to the copulatory activity of male and female rats. It appears probable that thyroidectomy is followed by reduction of general activity and lowered responsiveness to external stimulation in male and female alike; but because hers is the more passive mating rôle the female's sex reactions may be adequate for reproduction despite such loss; whereas in the male the depressing effects of the hormonal deficiency obviate the occurrence of sexual arousal which is essential to his copulatory performance.

A pre-puberally thyroidectomized bull is reported to have shown complete lack of mating behavior in adulthood (233). Ejaculates obtained from this animal were used successfully in artificially inseminating estrous cows, and this indicates that the spermatogenic functions of the testes were normal. Sexual arousal and copulatory behavior followed the administration of desiccated thyroid, dinitrophenol or testosterone propionate and the variety of the effective substances suggests that initial lack of sexual excitability was merely one reflection of a generalized metabolic disturbance.

4. *Thymus*. Presently available evidence give us little reason to believe that

the thymus has anything to do with sexual behavior. Removal of this body during infancy has no effect upon the breeding behavior of male or female rats during adult life (237). Although one experimenter has concluded that administration of thymocrecin reduces sexual responsiveness in the estrous female rat, the published report includes no quantitative measures of sexual receptivity, and both experimental and control groups were very small (70).

F. *Locus and Nature of Hormonal Action.* Evidence concerning the ways in which glandular secretions bring about changes in behavior is fragmentary and in most cases inferential, but a brief exposition of the principal recognized and suspected types of hormonal influence will serve to illustrate the problems involved and may suggest potentially fruitful areas for future investigation.

By virtue of their systemic effects products of several endocrine glands exert indirect control over the behavior of organisms. For example, because of their profound influence upon metabolic processes, secretions of the thyroid are essential to any and all behavioral responses which depend upon normal metabolic function. In the present discussion, however, such generalized glandular effects will be taken for granted, and attention will be confined to the more specific influences exerted by testicular and ovarian hormones upon courtship and mating.

1. *Loci of action.* a. *Extra-neural effects.* Gonadal hormones stimulate growth and development in various secondary sex structures which subsequently are employed in the execution of behavioral responses. Such effects are particularly evident in sub-mammalian forms. The advent of breeding coloration and gonopod development in certain fishes, hypertrophy of nuptial pads and vocal apparatus in male salentia, extrusion of "courtship teeth" in salamandria, and development of the copulatory organ in male birds of the order *Ansiformes* depend upon secretions from the sex glands (49).

Though usually less prominent, the secondary sex characters of mammals are in many cases under the influence of testicular or ovarian hormones. The final stages of antler growth in the male Virginia deer depend upon testicular androgen (319), and during the rutting season the buck employs these hormonally-stimulated structures in his battles with other adult males. The outstanding example of specific growth effects produced in mammals by testicular and ovarian hormones is seen in the stimulation exerted by these glandular products upon the accessory reproductive structures; and some authors seem inclined to believe that it is by this means exclusively that gonadal secretions affect behavior. Thus, one writer suggests that the most probable rôle of the testis hormone in the sex behavior of the rat is that of "fostering normal growth and development of the reproductive organs and perhaps maintaining a general 'sex tonus' " (314).

Such an extreme view is negated by evidence presented above to prove that sexual reactions may occur in the absence of the accessory reproductive tract. However, it is true that in the intact animal sensations arising from genital stimulation play an important part in the generation of intense sexual excitement. Furthermore it is known that gonadal hormones are essential to normal growth

of the copulatory organ. In the male rat, for example, penis growth is markedly inhibited by castration at birth. Animals subjected to this operation respond to androgen treatment in adulthood with the execution of numerous incomplete copulations; but because of their infantile phallic organs such individuals rarely achieve intromission and cannot attain the peak of arousal essential to ejaculation (50). Here then is an illustration of the indirect manner in which hormones may affect behavior by controlling the development of an essential effector structure, and it seems highly probable that additional examples of this type of effect will be demonstrated by future experimentation.

b. *Effects within the nervous system.* Gonadal hormones appear to exert the major part of their behavioral effect by way of the nervous system, although presently-available evidence does not permit any definitive statement as to the exact locus of hormonal action. It is conceivable that the sex gland secretions might influence behavior by altering the functional capacity of sensory receptors, but data to support such an assumption are lacking,[19] and the most probable loci of hormonal effects appear to lie within the brain and spinal cord.

Some experimentalists have concluded that estrogen controls mating responses in the female cat by virtue of its effect upon short-arc reflexes in the spinal cord (197); whereas other authorities prefer to believe that the influence of the hormone is exerted principally upon the integrative mechanisms of the diencephalon (24). Until our understanding of the neural basis for sexual behavior is more nearly complete, and until there is convincing evidence to indicate the limits of hormonal action, it would appear safest to assume that gonadal hormones affect all of the nervous units involved in mating activity.[20]

It has been explained that erection of the penis is mediated by low spinal circuits (23, 24), and there is reason to believe that androgen directly affects these primary nervous mechanisms. For example, completely decorticate male rats injected with large amounts of testosterone propionate exhibit constant erection, although they give no indication of sexual excitability and are totally unresponsive to the female in heat (46). Similar hormone treatment of the human male may be followed by priapism even in cases in which there is no increase in sexual desire (91).

It is clear that reproductive secretions also affect supraspinal levels, for studies of estrogenic effects in decerebrate female mammals indicate that ovarian hormones alter the functional capacity of midbrain mechanisms involved in the integration of the separate motor responses of copulation (24, 94).

[19] It has been reported that a condition of olfactory hyper-acuity occurs in women twenty-four to forty-eight hours before the onset of menstruation and persists until several days after the cessation of flow (107). Whether or not this altered sensitivity has any hormonal basis it is impossible to decide, but the possibility is evident.

[20] It seems probable, however, that sensitivity to hormonal facilitation varies from one sexual mechanism to another. It is known that different morphological sex characters require different amounts of gonadal hormone for normal development (235), and it is quite likely that similar differences exist between various behavioral units, with the result that a concentration of hormone sufficient to support one element in the total mating pattern may be too low to affect certain other segments of the sexual response.

Finally, it is to be noted that gonadal hormones influence activity within still other neural structures which contribute to sexual function without mediating any of the overt reactions comprising courtship or coition. When intact male rats are injected with androgen their responses to the estrous female are not materially altered. The promptness with which mating is initiated, the frequency of copulation, and the rapidity with which ejaculation occurs remain essentially unchanged. However, the hormone does exert a profound effect upon susceptibility to sexual arousal which is reflected in the increased variety of stimulus animals capable of evoking mating attempts on the part of hormonally-altered individuals. Thus, although untreated males are likely to mate only with the receptive female, the administration of testosterone propionate is followed by repeated and persistent attempts to copulate with diestrous female rats, anesthetized females, other males, and even with immature female guinea pigs (34). The tendency to become sexually aroused is markedly increased by androgen although the nature of copulatory reactions is not changed.

It seems probable that in humans the action of gonadal hormones upon higher nervous mechanisms may be at least partially responsible for the marked increase in erotic dreams described by castrates during periods of androgen treatment (91),[31] and for the increased pleasure in kissing and embracing members of the opposite sex which is often experienced under therapy (212).

2. *Nature of effect.* a. *Alteration of thresholds.* At the present state of knowledge theories concerning the probable nature of hormonal effects upon nervous tissue are highly speculative, but the more widely accepted suggestions deserve at least some consideration. Proposals that gonadal hormones "eroticise" the nervous system (71, 282), or "activate" (17) or "energize" (54) nervous patterns are too vague for objective analysis; but it is possible to deal critically with more explicit suggestions such as those to the effect that endocrine secretions actively stimulate nervous discharge in various circuits or centers (90, 249), or inhibit neural activity within one region or another (111, 168).

In the main, theories of the latter type receive only limited support, and the majority of writers on this subject have suggested that hormones neither stimulate nor inhibit nervous processes, but instead act as "sensitizers," lowering the threshold of particular neural elements to afferent stimulation (23, 40, 187, 230, 253, 265, 331). Direct support for hypotheses of this nature is lacking although some suggestive data are available. It is known, for instance, that in cases of adrenal insufficiency the irritability of the sympathetic system is drastically reduced (159).

b. *Basis for selective action.* Even if it were possible to demonstrate conclusively that one hormone increases the stimulability of nervous tissue whereas another fails to do so, there would remain the necessity of explaining the relatively specific behavioral effects exerted by secretions from different glands. Any adequate theory would have to account for the fact that in either males or fe-

[31] Although part of this psychic change may be traceable to increased genital activity during sleep.

males androgen tends to enhance masculine sexual reactions, while estrogen affects primarily the feminine mating pattern.

In essence this means that a particular biochemical agent exerts a strong effect upon one nervous mechanism and has at most only a mild influence upon a second one; while a different chemical substance is without effect upon the first mechanism but plays intensely upon the other neural circuit. Neuro-chemical relationships of this general type are well known in physiology. Various analgesics such as cobra venom reduce the sensitivity of nerve endings for pain without affecting other types of cutaneous receptor organs (195). Presumably the differential response of various nervous elements to a given compound is due to biochemical differences within the nervous structures themselves, and if this is the case it provides an explanation for the relative specificity of hormonal effects upon behavior.

It is possible that the neural circuits and centers involved in the mediation of feminine copulatory responses differ not only structurally but also biochemically from those responsible for masculine sexual reactions. Furthermore, the mechanisms concerned with feminine coital performance may be, by reason of their characteristic biochemical properties, especially sensitive to estrogen; whereas the mechanisms for masculine responses, having a different biochemical constitution, may be assumed to possess an affinity for androgenic compounds.

When the castrated female is injected with estrogen the hormone reacts with the nervous elements contributing to feminine mating behavior in such a manner as to increase their sensitivity to afferent stimulation, and sexually receptive responses appear. In similar fashion androgen increases the sensitivity of the mechanisms for masculine behavior in the castrated male. At the same time both sexes possess central nervous arrangements capable of mediating mating responses typical of the opposite sex; and therefore when androgen is injected in the genetic female, or when estrogen is administered to the male, the particular nervous mechanism which is sensitive to that chemical substance reacts promptly and the corresponding sexual pattern appears under conditions of appropriate external stimulation.

SUMMARY

The evidence reviewed in foregoing pages serves first to emphasize the pressing need for more intensive investigation of the physiology and psychology of sexual behavior. Existing data are too fragmentary to support any final conclusions, but a recapitulation of the more important aspects of our subject can be formulated in general terms providing it is explicitly understood that all such generalizations are extremely tentative and may be subject to numerous exceptions. As a matter of fact, recognizing the many lacunae in present-day knowledge, and realizing the prevalence of contradictory findings, no author could hope that his interpretations would prove to be more than rough approximations of the truth.

A. *Neurological Evidence.* 1. Many of the discrete sensory-motor components of the complete copulatory pattern are mediated by nervous circuits lying within

the spinal cord. This holds true for males and females and for all mammalian species which have been studied. Various genital reflexes and certain gross postural adjustments characteristic of coital activity survive and can be experimentally elicited in chronic spinal preparations. In the normal animal these spinal mechanisms are of course subject to facilitative and inhibitive influences from higher nervous levels.

2. Integration of the separate response units into a biologically effective behavior pattern is achieved by supra-spinal nervous elements. In the case of female rodents, lagomorphs and carnivores this function is carried out by mechanisms lying above the cord and below the cerebrum. Certain relatively restricted hypothalamic regions appear to be intimately involved in this activity. In any event it is evident that decerebrate females of the above-mentioned orders are capable of fertile mating with normal males. This does not mean that the female's sexual behavior is unaffected by decerebration; nor does it indicate that the normal individual's coital performance is independent of cerebral influences. Evidence to the contrary is available.

3. Nothing is known concerning the identity of sub-cortical, supra-spinal mechanisms which may contribute to the integration of the male's sexual responses.

4. In male rodents the cerebral cortex appears to be indispensable to sexual arousal, although so far as is known this part of the brain makes no essential contribution to the organization of the motor acts of copulation.

5. Studies of male cats indicate that extensive neo-pallial injury interferes with various sensory-motor adjustments involved in coitus; and this undoubtedly is due to the fact that in carnivores the motor and sensory cortices have assumed a larger rôle in the control of all types of voluntary activity. The totally decorticate male cat shows no sexual behavior, and gives little or no evidence of excitement when placed with a receptive female. The absence of sexual response in the decorticate male cat stands in striking constrast to the survival of copulatory performance in the decerebrate female of this species.

6. The extensive dependence of finer sensory-motor co-ordinations upon cortical centers in the primates makes it appear probable that in the monkey, ape, or human large scale injury to the neo-pallium would render copulation exceedingly difficult if not impossible. However, experimental proof for this assumption is lacking.

7. It is known that stimulation of particular cutaneous receptors, located in the genitals and elsewhere, contributes in major fashion to the intensification of sexual excitement; but the stimuli involved in pre-contact arousal are less well understood.

8. There is little direct evidence to reveal the sensory basis for sexual arousal in female mammals. It is established that females of several sub-primate species are capable of fertile copulation after combined removal of the end-organs for vision, olfaction and audition; but it seems quite probable that a broader approach to the problem would reveal deficits in the total mating pattern of such drastically desensitized individuals. All that has been demonstrated is that

they respond appropriately to the stimulation provided by direct bodily contact with a sexually active male.

9. Sexual arousal in males of lower mammalian species appears to depend, not upon a single, critical type of sensory cue (such as odor of the female in heat), but upon a pattern of stimulation involving simultaneous and repetitive activation of several receptor systems. No single type of end-organ is essential providing others remain.

10. Incidental observations suggest that in male and female primates the arousal of complete sexual excitement ordinarily rests upon stimulation of most of the major receptor structures.

B. *Behavioral Evidence.* Laboratory experiments dealing exclusively with behavior, and non-experimental observations reported by clinicians, commercial animal breeders and others provide a heterogeneous body of data much of which supplements the neurological findings summarized above.

1. In many animal species there is a close behavioral relationship between sexual arousal and various other emotions. Sexual responses tend to appear in emotionally-charged but non-sexual situations, and a variety of types of excitement may lead to active sexual behavior. Conversely, susceptibility to normal sexual arousal is subject to inhibition by fear, anxiety, etc. Findings of this nature indicate a functional relationship between the central nervous mechanisms for sexual activity and other neural units involved in non-sexual emotions.

2. A series of inter-sexual differences between male and female mammals of several sub-primate species provides evidence which supports certain conclusions drawn from neurological investigations.

(a) The sexual arousal of the male tends to be inhibited in a strange environment, or in a familiar environment which has previously been associated with painful stimulation; but comparable inhibition occurs much less frequently in the female. An environmental context in which heterosexual copulation has been experienced exerts a facilitative effect upon sexual responsiveness in the male, but does not do so in the female. Once arousal has occurred the male becomes relatively unresponsive to extraneous, non-sexual environmental stimuli; whereas the female is capable of carrying out her sexual rôle and simultaneously attending to other aspects of the total situation.

(b) Among rodents, carnivores and ungulates the male begins to show incomplete copulatory responses long before puberty; while the female displays none of the adult feminine sexual pattern until the first estrus.

These and other differences between sub-primate males and females are thought to be due in large measure to the fact that the male's sexual behavior is intimately influenced by the cerebral cortex, while the mating performance of the female is less directly dependent upon the neo-pallium.

3. Certain inter-sexual differences and similarities in the mating behavior of male and female primates throw some light upon the nervous basis for reproductive life at higher levels of the mammalian scale.

(a) Male and female monkeys, apes, and humans indulge in extensive and

frequent pre-puberal sex play; and in young chimpanzees and children such behavior may proceed as far as heterosexual copulation with intromission and pelvic thrusts. This phyletic increase in precocious sexual manifestations is regarded as a product of the higher development of the cerebral cortex and the increasing importance of this brain region to the mating activity of both sexes.

(b) Reproductively mature but sexually inexperienced female chimpanzees which are placed with a breeding male during the period of fertility react to the consort's copulatory advances in biologically effective fashion, and coition usually is carried out without difficulty or delay. In contrast, sexually inexperienced adult males of this species seem unable to effect union with the receptive female at the first opportunity, and a long period of practice and learning is necessary before the male is capable of normal copulatory adjustments. This inter-sexual difference is taken to be a reflection of differential importance of higher nervous mechanisms in the coital activities of the two sexes.

4. Females of many primate and sub-primate species often mount other females that are in heat, executing pelvic thrusts and otherwise behaving in masculine fashion. Such females display no morphological abnormalities of the reproductive system, and the capacity for masculine copulatory performance is regarded as an integral element in the feminine behavioral repertoire. Males of several species respond to the sexual advances of other males with the display of feminine copulatory reactions. Individuals which exhibit temporary inversion of the mating rôle revert promptly to the coital pattern of their own sex when opportunity is offered. The evidence is assumed to indicate an underlying bisexuality of the genetically-determined neuro-muscular constitution in both sexes, although in all cases structural and functional factors impose definite limits upon the completeness of bisexual mating behavior.

5. Behavior involving auto-genital stimulation is common to male and female mammals. In sub-primate species this activity consists chiefly of manipulating and licking the external genitalia during pre- and inter-copulatory intervals; but in the primates manual and oral masturbation occurs in a variety of situations and may continue to the point of orgasm. In observations of this type it is possible to discern a biological or evolutionary basis for certain auto-erotic practices common to human beings.

C. *Hormonal Evidence.* There are several lines of evidence to prove that gonadal hormones exert specific and powerful effects upon sexual behavior in all mammalian forms.

1. In seasonally-breeding species the occurrence of intense and complete sexual performance coincides with periods of maximal secretory function in the testes and ovaries. In year-round breeders the male is continuously active sexually, while the female's periods of receptivity occur when mature follicles are present and the estrogen level is high. The relationship between ovarian condition and sexual behavior is clear-cut in females of lower mammalian species, but the female monkey shows some willingness to receive the male at times when the ovaries do not contain ripe follicles; and even greater emancipation from hormonal control is seen in the mating activity of the female chimpanzee, who

exhibits periodicity of responsiveness but may permit copulation at any stage of the ovarian cycle. The highest degree of freedom from ovarian direction of sex desire obtains in the human female in whom non-hormonal factors exert at least equally important effects. These observations suggest a progressive, phylogenetic decrease in the importance of ovarian secretions to sexual drive in female mammals, and it has been suggested that this change is directly related to the increasing contribution made by higher nervous mechanisms.

2. Removal of the ovaries promptly and permanently abolishes sexual behavior in females of sub-primate species, and administration of ovarian hormones restores normal mating responses. Gonadectomized chimpanzees continue to display some copulatory behavior after operation, although the surviving sexual activity is less frequent and intense than that exhibited by the intact individual during the period of maximal genital swelling. Administration of estrogen to spayed apes occasions sexual swelling and the restitution of full copulatory behavior. Ovariectomized human females may experience a reduction in sexual desire and capacity for complete coital response; but in many cases no post-operative change is noted, and in either event the effects of replacement therapy are not at all predictable. Here again are indications of an evolutionary decrease in the importance of gonadal hormones to feminine receptivity, a change which may be related to progressive increase in the control exerted by the cerebral cortex.

3. Gonadectomy in adult males of sub-primate species results in fairly rapid loss of the more complete forms of copulatory behavior, although in many individuals incomplete coital responses persist indefinitely. In all cases administration of androgen restores normal sexual behavior. Copulation with intromission survives indefinitely in some castrated male chimpanzees, but androgen must be administered if the ejaculatory reaction is to appear. Among castrated humans individual differences are pronounced, but for years after operation many men suffer little reduction in sexual ability. In the male as in the female it appears that advancing phyletic status may be associated with progressive reduction in the importance of gonadal hormones to full sexual performance, and this change is regarded as a function of increase in the reliance of sexual functions upon higher brain centers, particularly the cerebral cortex.

D. *Concluding Comments.* In the course of mammalian evolution several changes in the physiological basis for sexual behavior appear to have taken place. Increase in the size and complexity of the neo-cortex has been accompanied by progressive encephalization of sensory-motor functions including many of those involved in courtship and mating. Concomitantly the subcerebral mechanisms which originally were capable of mediating sexual responses have come to be more and more dependent upon facilitative impulses from functionally associated circuits lying higher in the nervous system. Progressive encephalization of various sexual activities has resulted in increasing variability and modifiability, both of the types of stimuli adequate to elicit sexual activity and of the overt forms of behavior by which sexual excitement may be expressed. Finally, development of increasing dependence upon facilitation from the neo-pallium has

in some measure freed the more primitive sexual mechanisms from strict control by gonadal hormones.

REFERENCES

(1a) De Alba, J. and S. A. Asdell. Estrous behavior and hormones in the cow. J. Comp. Psychol. 39: 119–124, 1946.

(1) Allen, W. F. Effect of ablating the frontal lobes, hippocampi, and occipito-parieto-temporal (excepting pyriform areas) lobes on positive and negative olfactory conditioned reflexes. Am. J. Physiol. 128: 754, 1940.

(2) Allen, W. F. Effect of destroying three localized cerebral cortical areas for sound on correct conditioned differential responses of the dog's foreleg. Am. J. Physiol. 144: 415, 1945.

(3) Altmann, M. Interrelations of the sex cycle and the behavior of the sow. J. Comp. Psychol. 31: 481, 1941.

(4) Anderson, E. E. Consistency of tests of copulatory frequency in the male albino rat. J. Comp. Psychol. 21: 447, 1936.

(5) Anonymous. The care and breeding of the golden hamster. Turtox Service Leaflet no. 53: 1, 1945.

(6) Aronson, L. R. and G. K. Noble. The sexual behavior of Anura. 2. Neural mechanisms controlling mating behavior in the leopard frog, *Rana pipiens*. Bull. Am. Mus. Nat. Hist. 86: 89, 1945.

(7) Asdell, S. A., J. de Alba and J. S. Roberts. The levels of ovarian hormones required to induce heat and other reactions in the ovariectomized cow. J. Animal Sci. 4: 277, 1945.

(8) Bacq, Z. M. Impotence of the male rodent after sympathetic denervation of the genital organs. Am. J. Physiol. 96: 321, 1931.

(9) Bacq, Z. M. The effect of sympathectomy on sexual functions, lactation and the maternal behavior of the albino rat. Am. J. Physiol. 99: 444, 1931–32.

(10) Bacq, Z. M. and L. Brouha. Reserches sur la physiologie du système nerveux autonome. II. Le comportement des organes génitaux après énervation sympathetique. Arch. int. Physiol. 35: 250, 1933.

(11) Ball, J. Sex behavior of the rat after removal of the uterus and vagina. J. Comp. Psychol. 18: 419, 1934.

(12) Ball, J. Normal sex behavior in the rat after total extirpation of the vasa deferentia. Anat. Rec. 58: 49, 1934.

(13) Ball, J. Sexual responsiveness in female monkeys after castration and subsequent estrin administration. Psychol. Bull. 33: 811, 1936.

(14) Ball, J. Sex activity of castrated male rats increased by estrin administration. J. Comp. Psychol. 24: 135, 1937.

(15) Ball, J. The effect of male hormone on the sex behavior of female rats. Psychol. Bull. 34: 725, 1937.

(16) Ball, J. A test for measuring sexual excitability in the female rat. Am. J. Psychol. 107: 698, 1937.

(17) Ball, J. Male and female mating behavior in prepubertally castrated male rats receiving estrogens. J. Comp. Psychol. 28: 273, 1939.

(18) Ball, J. The effect of testosterone on the sex behavior of female rats. J. Comp. Psychol. 29: 151, 1940.

(19) Ball, J. Effect of progesterone upon sexual excitability in the female monkey. Psychol. Bull. 38: 533, 1941.

(20) Ball, J. and C. G. Hartman. Sexual excitability as related to the menstrual cycle in the monkey. Am. J. Obst. and Gynec. 29: 117, 1935.

(21) Bard, P. The effects of denervation of the genitalia on the oestrual behavior of cats. Am. J. Physiol. 113: 5, 1935.

(22) BARD, P. Oestrual behavior in surviving decorticate cats. Am. J. Physiol. **116**: 4, 1936.
(23) BARD, P. Central nervous mechanisms for emotional behavior patterns in animals. Res. Publ. Assoc. Nerv. Ment. Dis. **19**: 190, 1939.
(24) BARD, P. The hypothalamus and sexual behavior. in The hypothalamus and central levels of autonomic function. Res. Publ. Assoc. Nerv. Ment. Dis. **20**: 551, 1940.
(25) BARD, P. AND D. McK. RIOCH. A study of four cats deprived of neocortex and additional portions of the forebrain. Johns Hopkins Hospital Bull. **60**: 73, 1937.
(26) BARKER, J. R. On sex-intergrade pigs: Their anatomy, genetics, and developmental physiology. Brit. J. Exper. Biol. **2**: 247, 1924–25.
(27) BASCOM, K. F. The interstitial cells of the gonads of cattle with special reference to their embryonic development and significance. Am. J. Anat. **31**: 223, 1923.
(28) BAYES, W. R. Report of Mayor's Committee for the Study of Sex Offenses. Published by the City of New York, 1930–39.
(29) BEACH, F. A. Sex reversals in the mating pattern of the rat. J. Genet. Psychol. **53**: 329, 1938.
(30) BEACH, F. A. The neural basis of innate behavior. III. Comparison of learning ability and instinctive behavior in the rat. J. Comp. Psychol. **28**: 225, 1939.
(31) BEACH, F. A. Effects of cortical lesions upon the copulatory behavior of male rats. J. Comp. Psychol. **29**: 193, 1940.
(32) BEACH, F. A. Copulatory behavior of male rats raised in isolation and subjected to partial decortication prior to the acquisition of sexual experience. J. Comp. Psychol. **31**: 457, 1941.
(33) BEACH, F. A. Female mating behavior shown by male rats after administration of testosterone propionate. Endocrinol. **29**: 409, 1941.
(34) BEACH, F. A. Effects of testosterone propionate upon the copulatory behavior of sexually inexperienced male rats. J. Comp. Psychol. **33**: 227, 1942.
(35) BEACH, F. A. Comparison of copulatory behavior of male rats raised in isolation, cohabitation, and segregation. J. Genet. Psychol. **60**: 121, 1942.
(36) BEACH, F. A. Execution of the complete masculine copulatory pattern by sexually receptive female rats. J. Genet Psychol. **60**: 137, 1942.
(37) BEACH, F. A. Male and female mating behavior in prepuberally castrated female rats treated with androgens. Endocrinol. **31**: 673, 1942.
(38) BEACH, F. A. Copulatory behavior in prepuberally castrated male rats and its modification by estrogen administration. Endocrinol. **31**: 679, 1942.
(39) BEACH, F. A. Importance of progesterone to induction of sexual receptivity in spayed female rats. Proc. Soc. Exper. Biol. and Med. **51**: 369, 1942.
(40) BEACH, F. A. Analysis of factors involved in the arousal, maintenance and manifestation of sexual excitement in male animals. Psychosomatic Med. **4**: 173, 1942.
(41) BEACH, F. A. Analysis of the stimuli adequate to elicit mating behavior in the sexually inexperienced male rat. J. Comp. Psychol. **33**: 227, 1942.
(42) BEACH, F. A. Central nervous mechanisms involved in the reproductive behavior of vertebrates. Psychol. Bull. **39**: 200, 1942.
(43) BEACH, F. A. Sexual behavior of prepuberal male and female rats treated with gonadal hormones. J. Comp. Psychol. **34**: 285, 1942.
(44) BEACH, F. A. Effects of injury to the cerebral cortex upon the display of masculine and feminine mating behavior by female rats. J. Comp. Psychol. **36**: 169, 1943.
(45) BEACH, F. A. Effects of injury to the cerebral cortex upon sexually receptive behavior in the female rat. Psychosomatic Med. **6**: 40, 1944.
(46) BEACH, F. A. Relative effects of androgen upon the mating behavior of male rats subjected to forebrain injury or castration. J. Exper. Zool. **97**: 249, 1944.
(47) BEACH, F. A. Bisexual mating behavior in the male rat: Effects of castration and hormone administration. Physiol. Zool. **18**: 390, 1945.

(48) BEACH, F. A. Hormonal induction of mating responses in a rat with congenital absence of gonadal tissue. Anat. Rec. **92**: 289, 1945.
(49) BEACH, F. A. Hormones and behavior. A survey of interrelationships between endocrine secretions and patterns of overt response. In press. Paul B. Hoeber, Inc., N. Y.
(50) BEACH, F. A. AND A. M. HOLZ. Mating behavior in male rats castrated at various ages and injected with androgen. J. Exper. Zool. **101**: 91, 1946.
(51) BEACH, F. A. AND P. RASQUIN. Masculine copulatory behavior in intact and castrated female rats. Endocrinol. **31**: 393, 1942.
(52) VON BECHTEREW, W. Die Funktionen der Nervencentra. Vol. III. Gustav Fischer, Jena, 1911.
(53) BEETZ, F. Uber die von den weiblichen Geschlechtswerkzeugen auslosbaren Empfindungsqualitaten, mit besonderer Berucksichtigung des Schmerzsinnes. Arch. Gynaek. **162**: 106, 1936.
(54) BERG, I. A. Development of behavior: The micturition pattern in the dog. J. Exper. Psychol. **34**: 343, 1944.
(55) BINGHAM, H. C. Sex development in apes. Comp. Psychol. Monogr. **5**: 1, 1928.
(56) BLANDAU, R. J. Is the copulation plug essential for the en masse transport of spermatozoa into the uterine cornua of the albino rat? Anat. Rec. **91**: 266, 1945.
(57) BLANDAU, R. J. On the factors involved in sperm transport through the cervix uteri of the albino rat. Am. J. Anat. **77**: 253, 1945.
(58) BOLING, J. L. AND R. J. BLANDAU. The estrogen-progesterone induction of mating responses in the spayed female rat. Endocrinol. **25**: 359, 1939.
(59) BOLING, J. L., R. J. BLANDAU, B. RUNDLETT AND W. C. YOUNG. Factors underlying the failure of cyclic mating behavior in the albino rat. Anat. Rec. **80**: 155, 1941.
(60) BOLING, J. L., R. J. BLANDAU, J. G. WILSON AND W. C. YOUNG. Post-parturitional heat responses of new born and adult guinea pigs. Data on parturition. Proc. Soc. Exper. Biol. and Med. **42**: 128, 1939.
(61) BOLING, J. L., W. C. YOUNG AND E. W. DEMPSEY. Miscellaneous experiments on the estrogen progesterone induction of heat in the spayed guinea pig. Endocrinol. **23**: 182, 1938.
(62) BOND, C. R. The golden hamster (*Cricetus auratus*): care, breeding, and growth. Physiol. Zool. **18**: 52, 1945.
(63) BRADFORD, F. K. The auriculo-genital reflex in cats. Quart. J. Exper. Physiol. **27**: 271, 1938.
(63a) BRAMBELL, F. W. R. The reproduction of the wild rabbit *Oryctolagus cuniculus* (L.). Proc. Zool. Soc. Lond. **114**: 1–45, 1944.
(64) BROOKHART, J. M., F. L. DEY AND S. W. RANSON. Failure of ovarian hormones to cause mating reactions in spayed guinea pigs with hypothalamic lesions. Proc. Soc. Exper. Biol. and Med. **44**: 61, 1940.
(65) BROOKS, C. McC. The rôle of the cerebral cortex and of various sense organs in the excitation and execution of mating activity in the rabbit. Am. J. Physiol. **120**: 544, 1937.
(66) BROOKS, C. McC. A study of the mechanism whereby coitus excites the ovulation-producing activity of the rabbit's pituitary. Am. J. Physiol. **121**: 157, 1938.
(67) BROOKS, C. McC AND W. E. LEONARD. Effect of hypothalamic lesions on the estrous cycle of the rat. Progr. 25th Ann. meeting, Assoc. Study Internal Secretions, p. 13, 1941.
(68) BROSTER, L. R., C. ALLEN, H. W. C. VINES, J. PATTERSON, A. W. GREENWOOD, G. F. MARRIAN AND G. C. BUTLER. The adrenal cortex and intersexuality. Chapman & Hall, London, 1938.
(69) BROWN, T. G. Motor activation of the post-central gyrus. J. Physiol. **48**: 30, 1914.
(70) BRUNER, J. S. AND B. CUNNINGHAM. The effect of thymus extract on the sexual behavior of the female rat. J. Comp. Psychol. **27**: 69, 1939.

(71) Bullough, W. S. Endocrinological aspects of bird behaviour. Biol. Rev. **20:** 89, 1945.

(72) Burn, J. H. The relation of adrenaline to acetylcholine in the nervous system. Physiol. Rev. **25:** 377, 1945.

(73) Byron, C. S. and P. Katzen. Clinical effects of the oral use of methyl testosterone in eunuchoidism. J. Clin. Endocrinol. **1:** 359, 1941.

(74) Cannon, W. B. Bodily changes in pain, hunger, fear and rage. 2nd ed. D. Appleton-Century Co., New York, 1929.

(75) Carpenter, C. R. Sexual behavior of free ranging rhesus monkeys (*Macaca mulatta*). I. Specimens, procedures and behavioral characteristics of estrus. J. Comp. Psychol. **33:** 113, 1942.

(76) Carpenter, C. R. Sexual behavior of free ranging rhesus monkeys (*Macaca mulatta*). II. Periodicity of estrus, homosexual, autoerotic and non-conformist behavior. J. Comp. Psychol. **33:** 143, 1942.

(77) Ceni, C. Alienist and neurologist. Chap. 38, p. 359, 1917.

(78) Clark, G. Sexual behavior in rats with lesions in the anterior hypothalamus. Am. J. Physiol. **137:** 746, 1942.

(79) Clark, G. Effects of sex hormone therapy on a prepuberal male castrate chimpanzee. Federation Proc. **4:** 14, 1945.

(80) Clark, G. Statements followed by this reference number are supported by as yet unpublished data generously made available by Dr. George Clark and gathered by him in the course of his researches at the Yerkes Laboratories of Primate Biology at Orange Park, Florida.

(81) Clauberg, C. and K. W. Schultze. Die Folgen der Sterilisierung und der Kastration bei Mann und Frau. Ztschr. f. arztl. Fortbld. **31:** 425, 1934.

(82) Coe, W. R. Divergent pathways in sexual development. Science **91:** 175, 1940.

(83) Cole, H. H. Superfecundity in rats treated with mare gonadotropic hormone. Am. J. Physiol. **119:** 702, 1937.

(84) Cole, H. H., G. H. Hart and R. F. Miller. Studies on the hormonal control of estrous phenomena in the anestrous ewe. Endocrinol. **36:** 370, 1945.

(85) Cole, H. H. and R. F. Miller. Artificial induction of ovulation and oestrum in the ewe during anoestrum. Am. J. Physiol. **104:** 165, 1933.

(86) Commins, W. D. and C. P. Stone. Effects of castration on the behavior of mammals. Psychol. Bull. **29:** 493, 1932.

(87) Cooper, A. Observations on the structure and diseases of the testis. J. Churchill, London, 1830.

(88) Cooper, J. B. An exploratory study on African lions. Comp. Psychol. Monogr. **17:** 1, 1942.

(89) Crew, F. A. E. Abnormal sexuality in animals. III. Sex reversal. Quart. Rev. Biol. **2:** 427, 1927.

(90) Crisler, G., W. T. Booher, E. J. van Liere and J. C. Hall. The effect of feeding thyroid on the salivary conditioned reflex induced by morphine. Am. J. Physiol. **103:** 68, 1933.

(90a) Cvetkov, K. A trial application of prolan. Konevodstvo, **18**, 1941. (Taken from: Anim. Breed. Absts., **11:** 149, 1943.)

(91) Daniels, G. E. and E. S. Tauber. A dynamic approach to the study of replacement therapy in cases of castration. Am. J. Psychiat. **97:** 905, 1941.

(92) Davis, C. D. The effect of ablations of neocortex on mating, maternal behavior and the production of pseudopregnancy in the female rat and on copulatory activity in the male. Am. J. Physiol. **127:** 374, 1939.

(93) Davis, K. B. Factors in the sex life of twenty-two hundred women. Harpers, New York, 1929.

(93a) Day, F. T. Clinical and experimental observations on reproduction in the mare. J. Agric. Sci. **30:** 244–261, 1940.

(94) Dempsey, E. W. and D. McK. Rioch. The localization in the brain stem of the oestrous responses of the female guinea pig. J. Neurophysiol. **2:** 9, 1939.
(95) Dice, L. R. A family of dog-coyote hybrids. J. Mammal. **23:** 186, 1942.
(96) Dickinson, R. L. A thousand marriages: A medical study of sex adjustment. Williams & Wilkins, Baltimore, 1931.
(97) Diebschlag, E. Beobachtungen und Versuche an intakten und grosshirnlosen Eidechsen und Ringelnattern. Zool. Anz. **124:** 30, 1938.
(98) Dobzhansky, T. Genetics and the origin of species. Columbia University Press, New York, 1941.
(99) Dobzhansky, T. and E. Mayr. Experiments on sexual isolation in *Drosophila*. I. Geographic strains of *Drosophila willistoni*. Proc. Nat. Acad. Sci. **30:** 238, 1944.
(100) Dunn, C. W. Induced gynecomastia in the male. J. A. M. A. **115:** 2263, 1940.
(101) Durfee, T., M. W. Lerner and N. Kaplan. The artificial production of seminal ejaculation. Anat. Rec. **76:** 65, 1940.
(102) Dusser de Barenne, J. G. and V. D. Koskoff. Further observations on the flexor rigidity in the hind legs of the spinal cat. Am. J. Physiol. **107:** 441, 1934.
(103) Edwards, J. The effect of unilateral castration on spermatogenesis. Proc. Roy. Soc. B **128:** 407, 1940.
(104) Elder, J. H. Effects of theelin injections in normal prepubescent chimpanzees. Anat. Rec. **72:** 37, 1938.
(105) Elder, J. H. and R. M. Yerkes. The sexual cycle of the chimpanzee. Anat. Rec. **67:** 119, 1936.
(106) Ellis, A. The sexual psychology of human hermaphrodites. Psychosomatic Med. **7:** 108, 1945.
(107) Elsberg, C. A., E. D. Brewer and I. Levy. The sense of smell. IV. Concerning conditions which may temporarily alter normal olfactory acuity. Bull. Neur. Inst. N. Y. **4:** 31, 1935–36.
(108) Enders, R. K. Training the polygamous male. Fur Trade J. Can. **23:** 16, 70, 72, 1945.
(109) Enders, R. K. Reproduction in the mink. In press. J. Morph.
(110) Engle, E. T. The testis and hormones. Chap. 17 in Problems of Ageing, 2nd ed. Williams & Wilkins, Baltimore, 1942.
(111) Evans, L. T. Differential effects of ovarian hormones on territorial reaction time of the female *Anolis carolinensis*. Physiol. Zool. **10:** 456, 1937.
(112) Everett, J. W. The restoration of ovulatory cycles and corpus luteum formation in persistent-estrus rats by progesterone. Endocrinol. **27:** 681, 1940.
(113) Feiner, L. and T. Rothman. Study of a male castrate. J. A. M. A. **113:** 2144, 1939.
(114) Ferrier, D. The functions of the brain. Smith & Elder Co., London, 1876.
(115) Filler, W. and N. Drezner. Results of surgical castration in women over forty. Am. J. Obst. and Gynec. **47:** 122, 1944.
(116) Finesinger, J. E., J. V. Meigs and H. W. Sulkowitch. Clinical, psychiatric and psychoanalytic study of a case of male pseudohermaphroditism. Am. J. Obst. and Gynec. **44:** 310, 1942. (Psychol. Absts. **19** #1942, 1945.)
(117) Fisher, C., H. W. Magoun and S. W. Ranson. Systocia in diabetes insipidus. The relation of pituitary oxytocin to parturition. Am. J. Obst. and Gynec. **36:** 1, 1938.
(118) Folley, S. J. Experiments on the relation between the thyroid gland and lactation in the rat. J. Physiol. **93:** 401, 1938.
(119) Folley, S. J. and F. H. Malpress. The artificial induction of lactation in the bovine by the subcutaneous implantation of synthetic oestrogen tablets. J. Endocrinol. **4:** 1, 1944.
(120) Folley, S. J. and F. H. Malpress. Artificial induction of lactation in bovines by oral administration of synthetic oestrogens. J. Endocrinol. **4:** 23, 1944.

(121) Foote, R. M. Diethylstilbestrol in the management of psychopathological states in males. J. Nerv. and Ment. Dis. 99: 928, 1944.
(122) Ford, C. S. A comparative study of human reproduction. Yale University Publ. Anthrop., no. 32: 3, 1945.
(123) Frank, A. H. and R. M. Fraps. Induction of estrus in the ovariectomized golden hamster. Endocrinol. 37: 357, 1945.
(124) Frank, R. T. The female sex hormone. C. Thomas, Baltimore, 1929.
(125) de Fremery, P. and M. Tausk. Geschlechlicht activitat Kastrierter mannlicher Kanninchen nach Behandlung mit Testosterone Propionat. Acta brev. Neerl. Physiol. 7: 164, 1937.
(126) Friedgood, H. B. Induction of estrous behavior in anestrous cats with the follicle-stimulating and luteinizing hormones of the anterior pituitary gland. Am. J. Physiol. 126: 229, 1939.
(127) Friedgood, H. B. and A. B. Dawson. Physiological significance and morphology of the carmine cell in the cat's anterior pituitary. Endocrinol. 26: 1022, 1940.
(128) Fulton, J. F. Physiology of the nervous system. Oxford University Press, New York, 1938.
(129) Gantt, W. H. Experimental basis for neurotic behavior. Paul B. Hoeber, Inc., New York, 1944.
(130) Gaunt, R. and H. W. Hays. Rôle of progesterone and other hormones in survival of pseudopregnant adrenalectomized ferrets. Am. J. Physiol. 124: 767, 1938.
(131) Ghiselli, E. Encephalization of brightness discrimination in mammals. Science 86: 618, 1937.
(132) Girden, E. The acoustic mechanism of the cerebral cortex. Am. J. Psychol. 55: 518, 1942.
(133) Glass, S. J., H. J. Deuel and C. A. Wright. Sex hormone studies in male homosexuality. Endocrinol. 26: 590, 1940.
(134) Glass, S. J. and R. W. Johnson. Limitations and complications of organotherapy in male homosexuality. J. Clin. Endocrinol. 4: 540, 1944.
(135) Goldzieher, M. A. and M. S. Adler. Estrogen treatment in a female eunuchoid. J. Clin. Endocrinol. 1: 349, 1941.
(136) Goltz, F. Uber den Einfluss des Nervensystems auf die Vorgänge während der Schwangerschaft und des Gebärakts. Pflüger's Arch. 9: 552, 1874.
(137) Goltz, F. Über die Verrichtungen des Grosshirns. Bonn, Paris, 1881.
(138) Goltz, F. Der Hund ohne Grosshirn. Pflüger's Arch. 51: 570, 1892.
(139) Goodenough, F. L. Developmental psychology: An introduction to the study of human behavior. 2nd ed. D. Appleton-Century Co., New York, 1945.
(140) Gordon, M. B. Endocrine consideration of genito-urinary conditions in children. Urol. and Cutan. Rev. 45: 3, 1941.
(141) Gordon, M. B. and E. M. Fields. Observations on the effect of chorionic gonadotropin and male sex hormone on eunuchoidism. J. Clin. Endocrinol. 3: 589, 1943.
(142) Graller, D. L., H. Felson and L. Schiff. Use of stilbestrol in males. Progr. Association for the Study of Internal Secretions, p. 27, 1941.
(143) Gray, J. Aspects of animal locomotion. Proc. Roy. Soc. 128: 28, 1939.
(144) Green, W. W. and L. M. Winters. The effect of sex on the development of the pig. IV. Histological and endocrinological studies of the boar. J. Animal Sci. 4: 55, 1945. (Biol. Absts. 19: no. 8706, 1945.)
(145) Greenblatt, R. B. Hormonal factors in libido. J. Clin. Endocrinol. 3: 305, 1943.
(146) Greenblatt, R. B. Office endocrinology. 2nd ed. C. Thomas, Baltimore, 1944.
(147) Greenblatt, R. B., F. Mortara and R. Torpin. Sexual libido in the female. Am. J. Obst. and Gynec. 44: 658, 1942.
(148) Halverson, H. M. Genital and sphincter behavior of the male infant. J. Genet. Psychol. 56: 95, 1940.

(149) Hamilton, G. V. A study of sexual tendencies in monkeys and baboons. J. Animal Behav. 4: 295, 1914.
(150) Hamilton, J. B. Induction of penile erection by male hormone substances. Endocrinol. 21: 744, 1937.
(151) Hamilton, J. B. Evidences of marked stimulation by sex hormones in certain eunuchs, phenomena interpreted tentatively to result from changed function of the adrenal glands following castration. Anat. Rec. 85: 314, 1943.
(152) Hamilton, J. B. Demonstrated ability of penile erection in castrate men with markedly low titers of urinary androgens. Proc. Soc. Exper. Biol. and Med. 54: 309, 1943.
(153) Hammond, J. Recent scientific research on horse breeding problems. Yorkshire Agric. Soc. J., pp. 2–16, 1938.
(154) Hammond, J. Physiological aspects of bovine sterility. 57th Ann. Congr. Nat. Vet. Med. Assoc. Great Britain & Ireland, 1939.
(155) Hammond, J. Control of ovulation in farm animals. Nature 153: 702, 1944.
(156) Hammond, J. Induced ovulation and heat in anoestrous sheep. J. Endocrinol. 4: 169, 1945.
(157) Hammond, J. and F. T. Day. Oestrogen treatment of cattle: Induced lactation and other effects. J. Endocrinol. 4: 53, 1944.
(158) Hammond, J. and F. H. A. Marshall. Reproduction in the rabbit. Oliver & Boyd, London, 1925.
(159) Hartman, F. A. Functions of the adrenal cortex. Endocrinol. 30: 861, 1942.
(159a) Hathaway, S. R. Physiological psychology. D. Appleton-Century Co., New York. 1942.
(160) Head, H. The concept of nervous and mental energy. Brit. J. Psychol. 12: 126, 1923–24.
(161) Heller, C. G., R. E. Chandler and G. B. Myers. Effect of small and large doses of diethylstilbestrol on menopausal symptoms, vaginal smear and urinary gonadotropins in 23 oophorectomized women. J. Clin. Endocrinol. 4: 109, 1944.
(162) Heller, C. G., J. P. Farney and G. B. Myers. Development and correlation of menopausal symptoms, vaginal smear and urinary gonadotropin changes following castration in 27 women. J. Clin. Endocrinol. 4: 101, 1944.
(163) Hemmingsen, A. M. Studies on the oestrous-producing hormone (oestrin). Skand. Arch. f. Physiol. 65: 97, 1933.
(164) Herrick, C. J. Neurological foundations of behavior. Holt & Co., New York, 1924.
(164a) Herter, K. Die biologie der Europaeschen Igel. Monog. d. Wilsäuget. 5: 1–222, 1938.
(165) Hertz, R., R. K. Meyer and M. A. Spielman. The specificity of progesterone in inducing receptivity in the ovariectomized guinea pig. Endocrinol. 21: 533, 1937.
(166) Hines, M. The development and regression of reflexes, postures, and progression in the young macaque. Carnegie Inst. Wash. Publ. 541: 153, 1942.
(167) Hirschfeld, M. Sexual anomalies. Emerson Books, Inc., New York, 1944.
(168) Hoagland, H. On the mechanism of tonic immobility in vertebrates. J. Gen. Physiol. 11: 715, 1927–28.
(169) Hooker, C. W. The postnatal history and function of the interstitial cells of the testis of the bull. Am. J. Anat. 74: 1, 1944.
(170) Hu, C. K. and C. N. Frazier. Masculinization of adult female rabbit following injection of testosterone propionate. Proc. Soc. Exper. Biol. and Med. 42: 820, 1940.
(171) Ischlondsky, N. E. Hirnrinde und Psyche. Urban u. Schwarz, Wien, 1924.
(172) Ischlondsky, N. E. Facteurs determinant la constitution neuropsychique de l'enfant. Congr. Nat. de Psychiat. Infantile, Paris, 1937.

(173) Jacobsen, C. F. The effects of extirpations upon the higher brain processes. Physiol. Rev. 19: 303, 1939.

(174) Jenkins, M. The effect of segregation upon the sex behavior of the white rat as measured by the obstruction method. Genet. Psychol. Monogr. 3: 457, 1928.

(175) Kempf, E. J. The social and sexual behavior of infra-human primates with some comparable effects in human behavior. Psychoanalytic Rev. 4: 127, 1917.

(176) Kinsey, A. C. Statements followed by this reference number are based upon data generously made available by Dr. A. C. Kinsey of Indiana University, whose extensive interview study of sexual behavior in more than 10,000 humans is to be published in the future. 1946.

(177) Klüver, H. and P. C. Bucy. Preliminary analysis of functions of the temporal lobe in monkeys. Arch. Neur. and Psychiat. 42: 979, 1939.

(178) Kohler, W. The mentality of apes. Harcourt, Brace & Co., New York, 1925.

(179) Kokolsky, C. Practical mink breeding methods. Fur Trade J. Can. 22: 12, 28, 1945.

(180) Koster, R. Hormone factors in male behavior of the female rat. Endocrinol. 33: 337, 1943.

(181) Kunde, M., F. D'Amour, A. Carlson and R. Gustafson. The effect of estrin injections on the basal metabolism, uterine endometrium, lactation, mating, and maternal instincts of the adult dog. Am. J. Physiol. 95: 630, 1930.

(182) Lange, J. Die Folgen der Entmannung Erwachsener an des Hand der Kriegerserfahrungen dargestellt. Georg Thieme, Leipzig, 1934.

(183) Langworthy, O. R. Behavior disturbances related to decomposition of reflex activity caused by cerebral injury. An experimental study of the cat. J. Neuropath. and Exper. Neur. 3: 87, 1943.

(184) Lashley, K. S. Integrative functions of the cerebral cortex. Physiol. Rev. 13: 1, 1933.

(185) Lashley, K. S. Studies of cerebral function in learning. Comp. Psychol. Monogr. 11: 2, 1935.

(186) Lashley, K. S. Functional determinants of cerebral localization. Arch. Neur. and Psychiat. 38: 371, 1937.

(187) Lashley, K. S. Experimental analysis of instinctive behavior. Psychol. Rev. 45: 445, 1938.

(188) Lashley, K. S. Studies of cerebral function in learning. XII. Loss of the maze habit after occipital lesions in blind rats. J. Comp. Neur. 79: 431, 1943.

(189) Leathem, J. H. Experimental induction of estrus in the dog. Endocrinol. 22: 559, 1938.

(189a) Leathem, J. H. and J. A. Morrell. Induction of mating in the dog with pregnancy urine extract. Endocrinol. 23: 672, 1938.

(190) Levy, D. M. Fingersucking and accessory movements in early infancy. Am. J. Psychiat. 7: 881, 1928.

(191) Lipschutz, A. On the internal secretion of the sexual glands. J. Physiol. 51: 283, 1917.

(192) Lipschutz, A. On some fundamental laws of ovarian dynamics. Proc. Cambridge Phil. Soc., Biol. Sci. 2: 263, 1926–27.

(193) Louttit, C. M. Reproductive behavior of the guinea pig. II. The ontogenesis of the reproductive behavior pattern. J. Comp. Psychol. 9: 293, 1929.

(193a) Lowther, F. A study of the activities of a pair of *Galago senegalensis mohli* in captivity, including the births and postnatal development of twins. Zoologica 25: 435–462, 1940.

(194) Lurie, L. A. The endocrine factor in homosexuality. Report of treatment of 4 cases with androgen hormone. Am. J. Med. Sci. 208: 176, 1944.

(195) Macht, D. I. and M. B. Macht. Comparative effect of cobra venom and opiates on vision. J. Exper. Psychol. 25: 489, 1939.

(196) Macirone, C. and A. Walton. Fecundity of male rabbits as determined by "dummy matings." J. Agric. Sci. **28**: 122, 1938.
(197) Maes, J. P. Neural mechanism of sexual behavior in the female cat. Nature **144**: 598, 1939.
(198) Maes, J. P. Le mechanisme nerveux du comportement sexuel de la chatte. C. R. Soc. Biol. **133**: 95, 1940.
(199) Magnotti, T. L'importanza dell'olfatto sullo sviluppo e funzione delgi organ genitali. Boll. Mal. Oreich. **8**: 281, 1936.
(200) Makepeace, A. W., G. L. Weinstein and M. H. Friedman. The effect of progestin and progesterone on ovulation in the rabbit. Am. J. Physiol. **119**: 512, 1937.
(201) Malinowsky, B. The sexual life of savages in North-western Melanesia. Vols. 1 & 2. Horace Liveright, New York, 1929.
(202) Mallow, S. Beitrag zur Kastration von Sexualverbrechern. Ztschr. f. d. g. Neur. u. Psychiat. **148**: 501, 1933.
(203) Mansfeld, O. P. Eirstock und Geschlechtstrieb. Arch. f. Gynak. **117**: 294, 1922.
(204) Markley, M. H. and C. F. Bassett. Habits of a captive marten. Am. Midl. Nat. **28**: 604, 1942.
(205) Marquis, D. P. Can conditioned responses be established in the new born infant? J. Genet. Psychol. **39**: 479, 1931.
(206) Marshall, F. H. A. and J. Hammond. Experimental control by hormone action of the oestrous cycle in the ferret. J. Endocrinol. **4**: 159, 1945.
(206a) Marshall, F. H. A. and J. Hammond. Fertility and animal breeding. Minist. Agric. and Fish. Bull. No. 39: 1–44, 1945.
(207) Martino, G. Sul mecanismo dei riflessi genitali conditionati nei cane. Atti Accad. Lincei **29**: 695, 1939. (Psychol. Absts. **19** #1440, 1945.)
(208) Masserman, J. H. Principles of dynamic psychiatry. W. B. Saunders Co., Philadelphia, 1946.
(209) Mayr, E. Systematics and the origin of species. Columbia University Press, New York, 1942.
(210) McCullagh, E. P., R. McCullagh and N. F. Hicken. Diagnosis and treatment of hypogonadism in the male. Endocrinol. **17**: 49, 1933.
(211) McKenzie, F. F. and C. E. Terrill. Estrus, ovulation, and related phenomena in the ewe. Mo. Agric. Exp. Sta. Res. Bull. no. **264**: 1, 1937.
(212) Miller, N. E., G. Hubert and J. B. Hamilton. Mental and behavioral changes following male hormone treatment of adult castration, hypogonadism and psychic impotence. Proc. Soc. Exper. Biol. and Med. **38**: 538, 1938.
(213) Moehlig, R. C. Castration in the male. Notes on the hypothalamic pituitary gonadal system. Endocrinol. **27**: 743, 1940.
(214) Moelk, M. Vocalizing in the house-cat; A phonetic and functional study. Am. J. Psychol. **57**: 184, 1944.
(215) Moore, C. R. Sex gland transplantation and the modifying effect in rats and guinea pigs. Anat. Rec. **20**: 194, 1920.
(216) Moore, C. R. The physiology of the testis and application of male sex hormone. J. Urol. **47**: 31, 1942.
(217) Moore, C. R. Comparative biology of testicular and ovarian hormones. Biol. Symposia **9**: 3, 1942.
(218) Moore, C. R. and T. F. Gallagher. Seminal vesicle and prostate function as a testis-hormone indicator; the electric ejaculation test. Am. J. Anat. **45**: 39, 1930.
(219) Moore, C. R. and D. Price. Some effects of testosterone and testosterone propionate in the rat. Anat. Rec. **71**: 59, 1938.
(220) Moore, T. V. Physiological factors in the treatment of mental disorders. Psychiat. Quart. **16**: 765, 1942.
(221) Müller, L. R. and W. Dahl. Die Innervierung der Mannlichen Geschlechtsorgane. Deutsch. Arch. f. Klin. Med. **107**: 113, 1912.

(222) MULLIGAN, R. M. Feminization in male dogs. A syndrome associated with carcinoma of the testis and mimicked by the administration of estrogens. Am. J. Path. **20**: 865, 1944.
(223) MURIE, O. J. Notes on the sea otter. J. Mammal. **21**: 119, 1940.
(224) MURPHY, J. P. AND E. GELLHORN. Hypothalamic facilitation of the motor cortex. Federation Proc. **4**: 53, 1945.
(225) NELSON, W. O. AND T. F. GALLAGHER. Some effects of androgenic substances in the rat. A. The effect of male hormone extracts on the testes of hypophysectomized rats. Science **84**: 230, 1936.
(226) NISSEN, H. W. The effects of gonadectomy, vasotomy, and injections of placental and orchic extracts on the sex behavior of the white rat. Genet. Psychol. Monogr. **5**: 451, 1929.
(227) NOBLE, G. K. Neural basis of social behavior in vertebrates. Coll. Net. **14**: 121, 1939.
(228) NOBLE, G. K. AND L. R. ARONSON. The sexual behavior of Anura. 1. The normal mating pattern of *Rana pipiens*. Bull. Am. Mus. Nat. Hist. **80**: 127, 1942.
(229) NOBLE, G. K. AND R. BORNE. The effect of forebrain lesions on the sexual and fighting behavior in *Betta splendens* and other fishes. Anat. Rec. **79**, Suppl.: 49, 1941.
(230) NOBLE, G. K. AND B. GREENBERG. Testosterone propionate, a bisexual hormone in the American chameleon. Proc. Soc. Exper. Biol. and Med. **44**: 460, 1940.
(231) OESTING, R. B. AND B. WEBSTER. The sex hormone secretion of children. Endocrinol. **22**: 307, 1938.
(232) PEARSON, O. P. Reproduction in the shrew (*Blarina brevicauda* Say). Am. J. Anat. **75**: 39, 1944.
(233) PETERSON, W. E., A. SPIELMAN, B. S. POMEROY AND W. L. BOYD. Effect of thyroidectomy upon sexual behavior in the male bovine. Proc. Soc. Exper. Biol. and Med. **46**: 16, 1941.
(234) PETZSCH, H. Beiträge zur Biologie, insbesondere Fortpflanzungsbiologie des Hamsters (*Cricetus cricetus* L.). Kleintier u. Peltztier **12**: 11, 1936.
(235) PEZARD, A. La nôtion des "Seuils differentiels"; sabase experimental, son importance en endocrinologie sexuell. Rev. Fran. Endocrinol. **5**: 233, 1927.
(235a) PHILLIPS, R. W., R. M. FRAPS AND A. H. FRANKS. Hormonal stimulation of estrus and ovulation in sheep and goats. A review. Am. J. Vet. Res. **6**: 165–179, 1945.
(236) PIERON, H. Le Cerveau et Pensée. Bonn, Paris, 1923.
(237) PLAGGE, J. C. The thymus gland in relation to sex hormones and reproductive processes in the albino rat. J. Morph. **68**: 519, 1941.
(238) POTTENGER, F. M. AND D. G. SIMONSEN. A male sex-stimulating and female sex-repressing fraction from the adrenal gland. Endocrinol. **22**: 197, 1938.
(239) POYNTER, H. Testis hormone secretion in the rat under conditions of vasectomy or isolation. Anat. Rec. **74**: 355, 1939.
(240) PRATT, J. P. Sex functions in man. Chap. XXIV in Sex and internal secretions. 2nd ed. Williams & Wilkins, Baltimore, 1939.
(241) PRATT, J. P. A personal note on methyl testosterone in hypogonadism. J. Clin. Endocrinol. **2**: 460, 1942.
(242) PRATT, L. W. Behavior of bipedal rats. Bull. Johns Hopkins Hospital **72**: 265, 1943.
(243) PRICE, D. AND E. ORTIZ. The relation of age to reactivity in the reproductive system of the rat. Endocrinol. **34**: 215, 1944.
(244) RADO, S. A critical examination of the concept of bisexuality. Psychosomatic Med. **2**: 459, 1940.
(245) RAMSEY, G. V. The sex information of younger boys. Am. J. Orthopsychiat. **13**: 347, 1943.
(246) RAMSEY, G. V. The sexual development of boys. Am. J. Psychol. **56**: 217, 1943.
(247) RAYNAUD, A. Comportement sexuelle des souris femelles intersexuées. C. R. Soc. Biol. Paris **127**: 993, 1938.

(247a) REED, C. A. The copulatory behavior of small mammals. J. Comp. Psychol. **39**: 185–206, 1946.

(247b) REED, C. A. AND R. REED. The copulatory behavior of the golden hamster. J. Comp. Psychol. **39**: 7–12, 1946.

(248) RICE, V. A. Breeding and improvement of farm animals. 3rd ed. McGraw-Hill Book Co., New York, 1942.

(249) RICHTER, C. P. Animal behavior and internal drives. Quart. Rev. Biol. **2**: 307, 1927.

(250) RICHTER, C. P. Hypophyseal control of behavior. Cold Spring Harbor Symposia Quant. Biol. **5**: 258, 1937.

(251) RICKEY, E. The thyroid influence on the behavior of the white rat. Comp. Psychol. Monogr. **2**: 1, 1925.

(251a) RIDDOCH, G. The reflex functions of the completely divided spinal cord in man, compared with those associated with less severe lesions. Brain **40**: 264–402, 1917.

(252) RING, J. R. The estrogen-progesterone induction of sexual receptivity in the spayed female mouse. Endocrinol. **34**: 269, 1944.

(253) RIOCH, D. McK. Certain aspects of behavior in decorticate cats. Psychiatry **1**: 339, 1938.

(254) ROGERS, F. T. Studies of the brain stem. VI. An experimental study of the corpus striatum in the pigeon as related to various instinctive types of behavior. J. Comp. Neur. **35**: 21, 1922.

(255) ROOT, W. S. AND P. BARD. Erection in the cat following removal of lumbo-sacral segments. Am. J. Physiol. **119**: 392, 1937.

(256) ROTH, W. E. Ethnological studies among the North-West-Central Queensland Aborigines. Queensland Agent-General Offices, London, 1897.

(257) ROWE, A. W. AND C. H. LAWRENCE. Studies of the endocrine glands. IV. The male and female gonads. Endocrinol. **12**: 591, 1928.

(258) SALMON, U. J. Rationale for androgen therapy in gynecology. J. Clin. Endocrinol. **1**: 162, 1942.

(259) SAND, K. Experiments on the internal secretion of the sexual glands, especially on experimental hermaphroditism. J. Physiol. **53**: 255, 1919–20.

(260) SAUL, L. Cited in reference 129, 1944.

(261) SCHÄFER, E. A. Textbook of physiology. Vol. II. Macmillan, New York, 1900.

(262) SCHRADER, M. E. G. Uber die Stellung des Grosshirns im Reflexmechanismus des centralen Nervensystems der Wirbelthiere. Arch. exp. Path. Pharmakol. **29**: 55, 1892.

(263) SEARS, R. R. Survey of objective studies of psychoanalytic concepts. Soc. Sci. Res. Council Bull. **51**: 1, 1943.

(264) SEMANS, J. H. AND O. R. LANGWORTHY. Observations on the neurophysiology of sexual function in the male cat. J. Urol. **40**: 836, 1938.

(265) SEWARD, J. P. Studies on the reproductive activities of the guinea pig. III. The effect of androgenic hormone on sex drive in males and females. J. Comp. Psychol. **30**: 435, 1940.

(265a) SHADLE, A. R. Copulation in the porcupine. J. Wildlife Manag. **10**: 159–162, 1946.

(265b) SHADLE, A. R., M. SMELZER AND M. METZ. The sex reactions of porcupines (*Erethizon d. dorsatum*) before and after copulation. J. Mammal. **27**: 116–121, 1946.

(266) SHAPIRO, H. A. Effect of testosterone propionate on mating. Nature **139**: 588, 1937.

(267) SHORR, E. The menopause. Bull. N. Y. Acad. Med. **16**: 453, 1940.

(268) SLONAKER, J. R. Pseudopregnancy in the albino rat. Am. J. Physiol. **89**: 406, 1929.

(269) SLONAKER, J. R. Sex-drive in rats. Am. J. Physiol. **112**: 176, 1935.

(270) SMELSER, G. K. The response of guinea pig mammary glands to injected sex hormones and ovarian grafts and its bearing on the problem of sex hormone antagonism. Physiol. Zool. **6**: 396, 1933.

(272) SMITH, P. E. AND E. T. ENGLE. Experimental evidence regarding the rôle of the anterior pituitary in the development and regulation of the genital system. Am. J. Anat. **40**: 159, 1927.

(273) SMITH, P. E. AND S. L. LEONARD. Pregnancy urine injections in hypophysectomized rats. Proc. Soc. Exper. Biol. and Med. **30**: 1246, 1933.

(274) SMITH, P. E. AND S. L. LEONARD. Mating reactions of hypophysectomized male rats treated with pregnancy urine extracts. Endocrinol. **21**: 1250, 1937.

(275) SOLLENBERGER, R. T. AND J. B. HAMILTON. The effect of testosterone propionate upon the sexual behavior of castrated male guinea pigs. J. Comp. Psychol. **28**: 81, 1939.

(276) SPRAGG, S. D. S. Morphine addiction in chimpanzees. Comp. Psychol. Monogr. **15**: 1, 1940.

(277) STEINACH, E. Untersuchungen zur vergleichenden Physiologie der Mannlichen Geschlechtsorgane. III. Uber den Geschlechtstrieb der vor und nach der Pubertat kastrierten Ratten und uber das Schicksal der akzessorischen Geschlechtsdrusen in Folge der Kastration. Pflüger's Arch. **56**: 304, 1894.

(278) STEINACH, E. Umstimmung des Geschlechtscharakters bei Saugetieren durch Austausch der Pubertatsdrusen. Zentralbl. Physiol. **25**: 723, 1911.

(279) STEINACH, E. Willkürliche Umwandlung von Säugetiermannchen in Tiere mit ausgeprägt weiblichen Geschlechtskaracteren und Weiblicher Psyche. Pflüger's Arch. **144**: 71, 1912.

(280) STEINACH, E. Feminierung von Männchen und Maskulierung von Weibchen. Zentralbl. Physiol. **27**: 717, 1913.

(281) STEINACH, E. Pubertätsdrüsen und Zwitterbildung. Arch. f. Entwicklungsmech. **42**: 307, 1916.

(282) STEINACH, E. Sex and life. Viking Press, New York, 1940.

(283) STEINACH, E. AND H. KUN. Die entwicklungmechanische Bedeutung der Hypophysis als Aktivator der Keimdrüseninkretion. Med. Clin. **24**: 524, 1928.

(284) STEINACH, E. AND H. KUN. Luteingewebe und männliche Geschlechtscharaktere. Pflüger's Arch. **227**: 266, 1931.

(285) STIER, E. Disturbances of sexual functions through head trauma. Deutsch. Med. Wchnschr., Jan. 28, 1938. (Rev. in: J. Nerv. & Ment. Dis. **88**: 714.)

(286) STONE, C. P. Congenital sexual behavior of young male albino rats. J. Comp. Psychol. **2**: 95, 1922.

(287) STONE, C. P. Further study of the sensory functions in the activation of sexual behavior in the young male albino rat. J. Comp. Psychol. **3**: 469, 1923.

(288) STONE, C. P. Experimental studies of two important factors underlying masculine sexual behavior: the nervous system and the internal secretion of the testis. J. Exper. Psychol. **6**: 84, 1923.

(289) STONE, C. P. A note on "feminine" behavior in adult male rats. Am. J. Physiol. **68**: 39, 1924.

(290) STONE, C. P. The effects of cerebral destruction on the sexual behavior of rabbits. I. The olfactory bulbs. Am. J. Physiol. **71**: 430, 1925.

(291) STONE, C. P. The retention of copulatory ability in male rats following castration. J. Comp. Psychol. **7**: 369, 1927.

(292) STONE, C. P. Sexual drive. Chap. XVIII in Sex and internal secretions. 1st ed. Williams & Wilkins, Baltimore, 1932.

(293) STONE, C. P. The retention of copulatory ability in male rabbits following castration. J. Genet. Psychol. **40**: 296, 1932.

(294) STONE, C. P. Activation of impotent male rats by injections of testosterone propionate. J. Comp. Psychol. **25**: 445, 1938.

(295) STONE, C. P. Sex drive. Chap. XXIII in Sex and internal secretions. 2nd ed. Williams & Wilkins, Baltimore, 1939.

(296) STONE, C. P. Copulatory activity in adult male rats following castration and injections of testosterone propionate. Endocrinol. **24**: 165, 1939.

(297) STONE, C. P. Physiological psychology. Ann. Rev. Physiol. **7**: 623, 1945.

(298) STONE, C. P. AND L. W. FERGUSON. Temporal relationships in the copulatory acts of adult male rats. J. Comp. Psychol. 30: 419, 1940.

(299) STOPES, M. C. Married love. A. C. Fifield, London, 1921.

(300) STRUTHERS, P. H. Breeding habits of the Canadian porcupine (*Erethizon dorsatum*). J. Mammal. 9: 300, 1928.

(301) SYMONDS, C. P. Concussion and contusion of the brain and their sequelae. Injury of the skull and spinal cord. S. Brock, Ed. Williams & Wilkins, Baltimore, 1940.

(302) TAUBER, E. S. Effects of castration upon the sexuality of the adult male. Psychosomatic Med. 2: 74, 1940.

(303) TAUBER, E. S. AND G. E. DANIELS. Sex hormones and psychic conflict—a case report. Psychosomatic Med. 3: 72, 1941.

(304) TEN CATE, J. Akustische und optische Reaktionen der Katzen nach teilweisen und totalen Extirpationen des Neopalliums. Arch. neerl. Physiol. 19: 191, 1934.

(305) TERMAN, L. M., C. C. MILES AND ASSISTANTS. Sex and personality. McGraw-Hill Book Co., New York, 1936.

(306) TERRY, G. C. AND T. A. C. RENNIE. Analysis of parergasia. Nervous and Mental Disease Monograph, New York, 1938.

(307) THOMPSON, D. W. The works of Aristotle translated into English. Vol. IV. Historia. Clarenden Press, Oxford, 1910.

(308) THOREK, M. Experimental investigation of the rôle of the Leydig, seminiferous and Sertoli cells and effects of testicular transplantation. Endocrinol. 8: 61, 1924.

(309) TINKLEPAUGH, O. L. Sex cycles and other cyclic phenomena in a chimpanzee during adolescence, maturity and pregnancy. J. Morph. 54: 521, 1933.

(310) TINKLEPAUGH, O. L. The self-mutilation of a male Macacus rhesus monkey. J. Mammal. 9: 293–300, 1928.

(311) TORREY, T. W. The development of the urinogenital system of the albino rat. II. The gonads. Am. J. Anat. 76: 375, 1945.

(312) VRUWINK, J. AND S. POPENOE. Postoperative changes in the libido following sterilization. *Am. J. Obst. and Gynec. 19: 72, 1930.*

(313) WARNER, L. H. A survey of sex behavior in the white rat by means of the obstruction method. Comp. Psychol. Monogr. 4: 1, 1927.

(314) WHEELER, D. R. The inhibitory effects of punishment: An experimental study of the white rat. Thesis, Ph.D., Harvard University, unpublished, 1933.

(315) WIESNER, B. P. AND L. MIRSKAIA. On the endocrine basis of mating in the mouse. Quart. J. Exper. Physiol. 20: 274, 1930.

(315a) WIGHT, H. M. Reproduction in the eastern skunk (*Mephitus mephitus nigra*). J. Mammal. 12: 42–47, 1931.

(316) WILSON, J. G. AND W. C. YOUNG. Sensitivity to estrogens studied by means of experimentally induced mating responses in the female guinea pig and rat. Endocrinol. 29: 779, 1941.

(317) WIMSATT, W. A. Notes on breeding behavior, pregnancy, and parturition in some vespertilionid bats of the eastern United States. J. Mammal. 26: 23, 1945.

(318) WINDLE, W. F. Induction of mating and ovulation in the cat with pregnancy urine and serum extracts. Endocrinol. 25: 365, 1939.

(319) WISLOCKI, G. B. Studies on growth of deer antlers: II. Seasonal changes in the male reproductive tract of the Virginia deer (*Odocoileus virginianus borealis*); with a discussion of the factors controlling the antler-gonad periodicity. Essays in Biology: 629–654. University of California Press, Berkeley & Los Angeles, 1943.

(320) WITSCHI, E. AND C. A. PFEIFFER. The hormonal control of oestrus, ovulation and mating in the female rat. Anat. Rec. 64: 85, 1935.

(321) WRIGHT, J. G. The clinical application of oestrogens and gonadotropins. Middlesex Vet. 3: 15, 1943. (Biol. Absts. 18: #5074, 1944.)

(322) YEAGER, J. F. AND S. C. MUNSON. Physiological evidence of a site of action of DDT in an insect. Science 102: 305, 1945.

(323) YERKES, R. M. The mind of a gorilla. Genet. Psychol. Monogr. **2**: 1, 1927.
(324) YERKES, R. M. A chimpanzee family. J. Genet. Psychol. **48**: 362, 1936.
(325) YERKES, R. M. Social dominance and sexual status in the chimpanzee. Quart. Rev. Biol. **14**: 115, 1939.
(326) YERKES, R. M. AND J. H. ELDER. Oestrus, receptivity, and mating in the chimpanzee. Comp. Psychol. Monogr. **13**: 1, 1936.
(327) YOUNG, W. C. Observations and experiments of mating behavior in female mammals. Quart. Rev. Biol. **16**: 135, 311, 1941.
(328) YOUNG, W. C., E. W. DEMPSEY, C. W. HAGQUIST AND J. L. BOLING. Sexual behavior and sexual receptivity in the female guinea pig. J. Comp. Psychol. **27**: 49, 1939.
(329) YOUNG, W. C., E. W. DEMPSEY, H. I. MYERS AND C. W. HAGQUIST. The ovarian condition and sexual behavior in the female guinea pig. Am. J. Anat. **63**: 457, 1938.
(330) YOUNG, W. C. AND W. D. ORBISON. Changes in selected features of behavior in pairs of oppositely sexed chimpanzees during the sexual cycle and after ovariectomy. J. Comp. Psychol. **37**: 107, 1944.
(331) YOUNG, W. C. AND B. RUNDLETT. The hormonal induction of homosexual behavior in the spayed female guinea pig. Psychosomatic Med. **1**: 449, 1939.
(332) ZAVADOVSKY, B. M. Effects of increased fecundity in swine induced with the aid of the SPM gonad stimulating hormone. C. R. (Doklady) Acad. Sci. URSS **47**: 225, 1945.
(333) ZINGG, R. M. Feral man and cases of isolation. Am. J. Psychol. **53**: 487, 1940.
(334) ZITRIN, A. AND F. A. BEACH. Induction of mating activity in male cats. Ann. N. Y. Acad. Sci. **46**: 42, 1945.
(335) ZUCKERMAN, S. The social life of monkeys and apes. Harcourt, Brace & Co., New York, 1932.

RECENT ADVANCES IN THE STUDY OF BIOLOGICAL COMPETITION BETWEEN STRUCTURALLY RELATED COMPOUNDS

D. W. WOOLLEY

The Rockefeller Institute, New York City

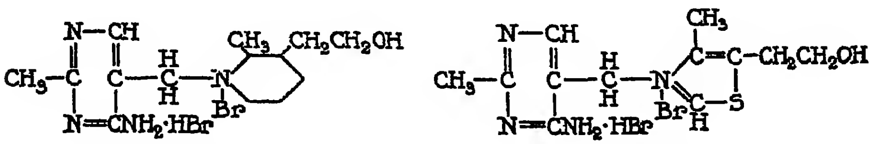
During the past few years compounds which are structurally similar to, but biologically antagonistic to metabolically important substances, have been synthesized in the laboratory or found to exist in nature. These agents which act in competition with vitamins, hormones, or other metabolites exhibit their effects by the production of signs of deficiency of the metabolite to which they bear structural resemblance. This situation has excited extensive study not only because of the light such a subtle relationship may throw on the mechanism of biochemical reactions, but also because some investigators have seen in it the possibility of using the findings as tools in pharmacology, biochemistry, or microbiology. Therefore, a review of existing knowledge of the phenomenon of competition between structurally related compounds may be advisable at this time.

Several excellent reviews dealing with this subject have appeared (17, 18, 3, 115, 159, 160) and these have recounted in some detail most of the experimental findings up to the summer of 1945. Therefore, no attempt will be made to reiterate that which has been said several times before. However, new information has come to light in the intervening time, and furthermore, sufficient delineation and integration have occurred to warrant further discussion.

Let us examine a typical case of the competition between a metabolite and its structural analog so that we may have a point of reference in attempting an understanding of the problems. The metabolite thiamine and its structural analog pyrithiamine provide an example in which many of the features common to the phenomenon may be clearly seen. The formulae of this pair of compounds are shown in figure 1. The two differ in that the thiazole ring of the vitamin is replaced by a pyridine ring in the pyrithiamine, that is, a S atom is exchanged for —CH:CH—. Pyrithiamine was first prepared by Tracy and Elderfield (161) in the anticipation that it would have thiamine potency, but in testing for such activity Robbins (121) found that for two species of thiamine-requiring fungi it was harmful rather than beneficial. Woolley and White (116), who were interested in the phenomenon of competition between structurally related compounds, then fed pyrithiamine to mice, and observed that signs characteristic of thiamine deficiency were produced. The deficiency syndrome appeared with greater speed and more pronounced intensity, and sharper definition of the typical manifestations of avitaminosis B_1 than had ever been achieved in mice by mere restriction of thiamine intake. Hyperirritability, unsteady gait, head retraction, convulsions (initiated by twirling the animal while it was held suspended by the tail head downward), prostration, loss of weight and death have been seen frequently in various animals receiving no thiamine in the diet, but usually only a few of these signs may be observed in any one species, and

with mice, usually only the latter two are noted. However, the entire sequence followed the administration of pyrithiamine. The basal diet fed to these animals contained adequate thiamine to maintain health in the absence of the drug, and when a further quantity of the vitamin was added, the action of the drug was prevented. Furthermore, administration of thiamine cured mice which had been affected by pyrithiamine.

Now, the toxicity of pyrithiamine (i.e., its ability to call forth signs of thiamine deficiency) depended on the ratio of pyrithiamine to thiamine in the diet, and not on the absolute amount of the drug. The effects of 100 lethal doses were completely prevented by increasing the thiamine content of the diet 100-fold.



Pyrithiamine Thiamine

Fig. 1. Structural formulae of thiamine and pyrithiamine

TABLE 1

Response of mice to various doses of thiamine and pyrithiamine

PYRITHIAMINE	THIAMINE	ANIMALS SHOWING DEFICIENCY SIGNS	AVERAGE CHANGE IN WEIGHT
gamma per day	*gamma per day*	*per cent*	*grams per week*
0	1.6	0	+3.0
600	1.6	100	−0.2
300	1.6	100	+1.9
100	1.6	100	+2.5
100	2.0	75	+2.4
50	2.0	0	+3.5
600	61.6	0	+3.1
2,000	60.0	0	+3.6

In other words, pyrithiamine and thiamine competed with each other over a wide range of concentration for the functions of the latter in the mouse, and were thus antagonistic structural analogs. Data to illustrate this point are shown in table 1.

The biological effects of pyrithiamine were not limited to the action on mice, for the drug proved to be capable of inhibiting the growth of a variety of yeasts, molds, and bacteria (117). In preventing the growth of such forms the same type of competition with thiamine was found as had been observed with mice. Since the microbiostasis could be prevented by increasing the thiamine content of the media, the action of the agent could be said to reside in its ability to cause thiamine deficiency of the organisms. In much of the work on the competition between structurally related compounds, bacteria and other microorganisms

have been used as test objects to demonstrate the antagonism. This has been due partly to the fact that such forms make possible more rapid, and usually more quantitative assays than those with laboratory animals. Screening of many substances for a given activity is thus facilitated.

Not all microorganisms were inhibited in growth by pyrithiamine. Only those species which needed thiamine (or its constituent pyrimidine and thiazole parts) as essential growth stimulants were susceptible to its action. The drug was thus a highly selective inhibitor of microbial growth since it distinguished between those organisms which required the vitamin and those which did not.

Competitive relationships between metabolites (vitamins, hormones, and other biologically active compounds) have been observed for many years. Perhaps the most outstanding observation was that of Woods (20) in 1940. He showed that the bacteriostatic action of sulfanilamide was reversed by the structurally related p-aminobenzoic acid. The relationship between the drug and the metabolite was competitive over a wide range of concentration, so that, just as we have seen above, the activity of the sulfanilamide depended more on the ratio of it and p-aminobenzoic acid than on the absolute amount present. At the time of Woods' work, p-aminobenzoic acid was not known to be a metabolite, but he showed that the natural inhibitor of sulfonamide action (which was present in yeast) had properties similar to those of this substance. Shortly after his postulation of the rôle of this compound in metabolism abundant proof was forthcoming from varied sources (162, 163) to substantiate his claim. Not only was p-aminobenzoic acid found in many living cells, but it was shown to be an essential nutrient for some microorganisms, and finally to be an integral part of the vitamin pteroyl glutamic acid (folic acid) (164).

As a result of his observations, Woods proposed the idea that sulfanilamide and its derivatives owed their bacteriostatic action to the competition which existed between them and the metabolite p-aminobenzoic acid. The bacteria were thus effectively deprived of the latter, and in such condition were unable to thrive within the host.

The discovery of the relationship between the sulfonamide drugs and p-aminobenzoic acid was not the first example to be observed of competition between structurally similar compounds, but it was one of the most important because it offered an explanation for the mode of action of a set of new and dramatic therapeutic agents. The practical importance of the drugs was sufficient to stimulate interest in the further investigation of a phenomenon (i.e., that of competition) which previously had been noted incidentally and somewhat casually (3, 115).

In looking back over the literature one frequently finds experimental observations which were baffling to their discoverers, and uncongenial to the modes of biochemical thought at the time, but which now seem understandable, or at least reconcilable in terms of the present knowledge of antagonism between structurally similar compounds. One interesting example of this is the claim of Kuhn (42) that the sex hormone of certain algae was not a single compound, but rather a manifestation of a ratio between two structurally allied carotenoids,

cis- and *trans*-dimethyl crocetin. Sex was determined by the ratio between these analogs. Specific biological effects had been attributed so universally to single and distinct chemical substances that even in 1940 when these results with algae were summarized they sounded somewhat heretical. Some investigators have come to feel now, however, in view of the demonstrable competition between androgen and estrogen in higher animals, and the undoubted structural similarity and coexistence of these hormones in nature, that sex even in higher animals may in part be an expression of competition between related structures (115).

Since 1940, antagonistic analogs have been sought with such industry that there is at least one for each of the water-soluble vitamins, and for two of the four fat-soluble ones, and for certain hormones, amino acids, porphyrins, and purines. Usually for each of these metabolites there is not one but several types of inhibitor, differing qualitatively in their biological effects. Because most of these cases have been reviewed adequately elsewhere (17, 18, 3, 115, 159, 160) nothing will be gained by reiteration. Nevertheless, for purposes of reference and of orientation Table 2 has been compiled to show which metabolites have been converted into antagonistic compounds by alteration of their structures, and the types of change which have been found effective. The table includes information published since the last review of the subject as well as the older examples.

General Considerations. (*a*) *Concerning reversal.* The surest indication that the antagonistic effects produced by a structural analog are due to the competition of that analog with its related metabolite is the ability of the latter to reverse the action of the agent. The evidence is strongest when over a considerable range of concentrations a constant ratio exists between the amounts of the two related compounds which just counterbalance each other. This ratio which is known as the inhibition index (117) is a constant for many of the cases listed in table 2. It is so, for example, with the sulfonamide drugs and p-aminobenzoic acid, and with pyrithiamine and thiamine. However, it does not hold universally. The situation with 3,3′-methylenebis-(4-hydroxy-coumarin) and vitamin K will serve to illustrate this point. With this pair of compounds it is possible to demonstrate (142, 165) that the pharmacological action of the former in bringing about the signs of vitamin K deficiency is reversed by sufficient amounts of the latter, but this is only true over a restricted range of concentration of the drug. As the quantity of the drug is increased, a point is soon reached beyond which it is not possible to achieve such reversal. Furthermore, there is no evidence to show that the inhibition index (117) is a constant even over the limited range where reversal can be demonstrated.

An additional obscurity in this case is the fact that very large amounts of vitamin K are required to antagonize the action of the drug. About 25 mgm. of the vitamin must be used to erase the effects of 10 mgm. of the drug in a rat. This means that the inhibition index is less than 1. For most inhibitor-metabolite pairs the index is much greater than 1.

Another example of this atypical behavior is to be found with 2,3-dichloro-

TABLE 2

Metabolites and structural analogs which have been found antagonistic to them

METABOLITE	ANALOG	STRUCTURAL ALTERATION	BIOLOGICAL SYSTEM AFFECTED	REFERENCES
Acetic acid	Fluoro-acetic aid	F for H	Acetate oxidase	158
Adenine	Benzmidazole and derivatives	2 C for 2 N; side chain alterations	Microorganisms, animals	1, 2, 3
	Triazolopyrimidines	N for C	Microorganisms	4
β-Alanine	β-Amino-butyric acid	CH_3 for H	Yeast	5
	Propionic acid	H for NH_2	Bacteria	157
	Asparagine	COOH for H; $CONH_2$ for COOH	Yeast	6
α-Alanine	Glycine	H for CH_3	Bacteria	7
Amino acids	α-Amino-sulfonic acids	SO_3H for COOH	Bacteria	8, 9
	Other amino acids		Bacteria	10, 7, 11
Aspartic acid	Hydroxy-aspartic acid	OH for H	Bacteria	12
Ascorbic acid	Glucoascorbic acid	Addition of CHOH and optical inversion	Animals; liver enzymes	13, 14, 15
p-Aminobenzoic acid	Sulfanilamide and derivatives	SO_2NH_2 or derivative for COOH	Microorganisms	16, 17, 18, 19, 20, 32
	p-Amino benzamide	$CONH_2$ for COOH	Bacteria	21, 22
	Carbarsone and related arsenicals	As for C in a COOH group; derivatives of this	Microorganisms, animals	21, 23, 24, 25, 26, 27, 28
	Phosphanilic acid	PO_3H_2 for COOH	Microorganisms	155
	Heterocyclic acids, e.g., 6-amino-nicotinic acid	N or S for C	Bacteria	29
	Ring-substituted PAB	Halogen or alkyl for H	Bacteria	30, 29, 31
	p-Amino-acetophenone and derivatives	COR for COOH	Bacteria	23, 33
	p-Nitrobenzoic acid	NO_2 for NH_2	Bacteria	154
Biotin	Desthiobiotin and derivatives	2 H for S	Microorganisms	34, 35, 36, 37
	Biotin sulfone	SO_2 for S	Microorganisms	35
	Ureylenecyclohexyl aliphatic acids	2 C for S; and derivatives with shorter side chains	Microorganisms	39, 40
	Desthioisobiotin	Loss of S, geometric isomerism	Insects	78
Choline	Triethyl choline	3 ethyls for 3 methyls	Frog muscle, mice	41
Cocarboxylase	Thiamine-thiazole pyrophosphate	Loss of pyrimidine portion	Carboxylase	125
Cis- or *trans-*crocetin dimethyl ester	*Trans-* or *cis-*crocetin dimethyl ester	Geometric isomerism	Algae	42
Folic acid	See Pteroyl glutamic acid			
Glutamic acid	Methionine sulfoxide	$SOCH_3$ for COOH	Bacteria	43, 44
Guanine	Triazolopyrimidines	N for C	Bacteria	4
	Benzimidazole	2 C for N	Microorganisms	2
Histamine	Imidazole and derivatives	Elimination or substitution of part of the molecule	Smooth muscle, histamine shock in animals	17, 45, 46, 47, 48, 49, 50, 51, 52

TABLE 2—*Continued*

METABOLITE	ANALOG	STRUCTURAL ALTERATION	BIOLOGICAL SYSTEM AFFECTED	REFERENCES
Hypoxanthine	Hydroxy-triazolo-pyrimidine	N for C	Bacteria	4
Indoleacetic acid	Phenyl butyric acid	Elimination of N and shift of one C	Plants	53
Inositol	Hexachlorocyclohexane	6 Cl for 6 OH	Yeast	54
Isoleucine	Leucine	Position isomerism of one CH_3	Bacteria	55
Isoleucine + valine	Phenyl alanine or nor-leucine or nor-valine	C_6H_5 for alkyl position isomerism	Fungi	56
Leucine	d-Leucine	Optical inversion	Bacteria	57
Lysine	Arginine	Guanidino for amino, elimination of CH_2	Fungi	58
Methionine	Methoxinine	O for S	Bacteria	4
	Ethionine	CH_3 for H	Bacteria, animals	59, 60
Nicotinic acid (or amide)	Pyridine-3-sulfonic acid (or amide)	SO_3H for COOH	Microorganisms, animals	61, 62, 63, 64, 65, 66, 67
	3-acetyl-pyridine	$COCH_3$ for COOH	Animals, not in microorganisms	33, 68
	5-thiazole carboxamide	S for CH=CH	Certain bacteria	65, 69
Pantothenic acid	Thiopanic acid (pantoyl taurine) and derivatives	SO_3H and derivatives for COOH	Microorganisms, pantothenate utilizing enzymes, not animals	70, 71, 72, 67, 73, 74, 75, 76, 77, 79, 80, 81, 82, 83, 6, 85, 86, 92, 93
	Pantothenyl alcohol	CH_2OH for COOH	Microorganisms, not animals	84, 87
	α- or β-Methylpantothenic acid	CH_3 for H	Microorganisms	88, 89, 202
	Substituted panto-amides other than the above	Alkyl or OH- and NH_2-alkyl for CH_2CH_2COOH	Microorganisms	90, 91
	Phenyl pantothenone	COC_6H_5 for COOH	Microorganisms	94, 95, 3, 96
	Salicylyl β-alanine	o-hydroxy-benzoyl for pantoyl	Microorganisms	97
	γ'-Methyl pantothenic acid	CH_3 for H	Bacteria	156
Phenyl alanine	β-Hydroxyphenyl alanine	OH for H	Bacteria	98
	Thienyl-alanine	S for CH=CH	Microorganisms	99, 98, 100
Porphyrins such as hematin and protoporphyrin	Porphyrins lacking vinyl groups		Bacteria	101
Pteroylglutamic acid	Pteroyl-triglutamic acid	Addition of two glutamic acids	Transplanted tumors	102
	Xanthopterin	Loss of p-amino-benzoyl glutamic acid	Transplanted tumors	102
Pyridoxine	Desoxypyridoxine	H for OH	Chickens	103
Riboflavin	6,7-dichloro-riboflavin	2 Cl for 2 CH_3	Microorganisms	104
	Isoriboflavin	Shift in position of CH_3	Animals, not bacteria	105, 110
	Corresponding phenazine	2 C for 2 N, 2 NH_2 for 2 OH	Microorganisms, animals	106, 109
	Galactoflavin	Dulcityl for ribityl	Animals	107
	Lumiflavin	CH_3 for ribityl	Bacteria	108, 109, 204
	Araboflavin	Inversion of position of OH	Animals	203

TABLE 2—*Concluded*

METABOLITE	ANALOG	STRUCTURAL ALTERATION	BIOLOGICAL SYSTEM AFFECTED	REFERENCES
Succinic acid	Malonic acid	Loss of CH_2	Succinic oxidase	111
	Sulfonated succinic acid	SO_3H for H	Succinic oxidase	112
Testosterone	Estradiol	Benzene ring for cyclohexane ring	Animals	113, 114, 115
Thiamine	Pyrithiamine	CH=CH for S	Animals, microorganisms	116, 117, 118, 119, 120, 121, 201
	Oxythiamine	OH for NH_2	Animals, fish thiaminase	122
	Butyl thiamine	Butyl for CH_3	Animals	123
	Aminobenzylmethyl-thiazolium chloride	2 C for 2 N, loss of side chains	Fish thiaminase	124
Thymine	5-substituted pyrimidines	OH or NO_2 or Br for CH_3	Bacteria	126
α-Tocopherol	α-Tocopherol quinone	Opening of ring by addition of H_2O	Animals	127
Thyroxine	Ethers of diiodotyrosine	p-nitro-benzyl or p-nitro-phenylethyl or benzyl for p-hydroxy-diiodophenyl	Tadpoles	128
Tryptophane	Indoleacrylic acid	Loss of NH_2	Bacteria	129
	Naphthylacrylic acid	Loss of NH_2, C=C for N	Bacteria	130
	Styrylacetic acid	Loss of NH_2, substitution of aliphatic unsaturated side chain for pyrrol ring	Bacteria	130
	5-methyl-tryptophane	CH_3 for H	Bacteria	131
Uracil	Barbituric acid	OH for H	Bacteria	132
	Thiouracil	S for O	Bacteria	133
Vitamin K	Dicoumarol and derivatives	O for C, side chain alterations	Animals	134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148
	Iodinin	2 N for 2 C, side chain alterations	Bacteria	149
	α-Tocopherol quinone	2 CH_3 for benzene ring	Animals	127
	2,3-dichloronaphthoquinone	2 Cl for alkyl side chain	Microorganisms	150, 151
	2-substituted-3-hydroxy-naphthoquinones	OH for H, change in alkyl substituent	Animals, not bacteria	152, 153

naphthoquinone and the structurally similar vitamin K. Here, too, competition can be demonstrated over only a narrow range of concentration and with only certain species (151). A plausible explanation of this case was given by the finding that the vitamin itself was harmful in large doses (151). However, there are examples of the failure to demonstrate typical competition where this explanation of toxicity of the metabolite cannot be applied.

Why, then, are such cases included in a discussion of competition between structurally related compounds? In the first place, several of these questionable

analogs call forth in animals many of the signs characteristically associated with a deficiency of the metabolite to which the drug is structurally related. For example, 3,3′-methylenebis-(4-hydroxy-coumarin) brings about subcutaneous hemorrhages, and prolonged blood clotting-time attributable to decrease in the plasma prothrombin. These are typical manifestations of avitaminosis K. Again, α-tocopherol quinone causes pregnant mice to fail to produce young in a manner similar to that seen in tocopherol deficiency (127). Hence, despite the failure to reverse these effects with the vitamin there is some reason to relate the action of the drug to an interference with the biological action of the metabolite. However, in the cases of mere inhibition of microbial growth by an analog of a metabolite, the response (i.e., inhibition of growth) is too nonspecific to link the action of the drug with that of the metabolite unless reversal of the effect of the former by the latter can be demonstrated.

In the second place, the ability to demonstrate competitive reversal frequently depends on the selection of the proper species or biological system. Take for example the classical case of the sulfonamide drugs and p-aminobenzoic acid. Although the competition of these substances in producing bacteriostasis has been demonstrated abundantly with a wide variety of species, Tamura (166) has shown that no reversal of sulfonamide action occurred with *B. tularensis*. Again, Woolley and Collyer found (94) that phenyl pantothenone prevented the growth of all microbial species examined, but that reversal of this effect could be achieved with the related metabolite pantothenic acid only in those forms which were stimulated to grow by the vitamin. With the non-pantothenic acid-requiring microorganisms the vitamin was unable to reverse the inhibition of growth. If only organisms not requiring the vitamin had been examined, what would have been concluded about the action of the drug?

In view of these aspects, it seems that while reversal of the action of a structural analog by increased amounts of the metabolite is the acid test, it is not an entirely sufficient one. In deciding which cases shall be accepted to membership in the phenomenon of competitive inhibition, and which shall be excluded, no single criterion will suffice. Just as with the definition of enzyme action (115), here, too, one must content himself with descriptions of features common to many individual cases but not necessarily shared by all. Nature is not a series of discrete quanta of phenomena, but rather a continuous gradation of them.

(*b*) *Dependence of antagonistic action on nutritional requirement.* Frequently an analog is found to be antagonistic to a metabolite only if the latter is a nutritional requirement of the organism with which the test is being conducted. For species which do not need the metabolite in their food, and can synthesize it themselves, the analog has no detectable effect. Thus thiopanic acid affects only certain bacteria which will not grow without pantothenic acid, and pyrithiamine retards the growth of microorganisms only if they demand an exogenous supply of thiamine or its constituent thiazole and pyrimidine parts (117). Many exceptions to this generalization are known. For example, the sulfonamide drugs, phenyl pantothenone and its derivatives, benzimidazole and its derivatives, and 2,3-dichloronaphthoquinone act against organisms irrespective

of their nutritional requirements for the related metabolites. Some antagonists, such as phenyl pantothenone or glucoascorbic acid (13, 14), exhibit properties midway between these extremes, for although they affect organisms without regard to nutritional requirements, their action is reversed by the related metabolite only when the particular organism needs the latter as a growth factor. The case of phenyl pantothenone was mentioned in this respect in section (*a*) above. That of glucoascorbic acid will illustrate a similar phenomenon in the animal kingdom. This substance produced a disease resembling scurvy in both mice and guinea pigs, but the condition was prevented by ascorbic acid only in the latter, or ascorbic acid-requiring, species, and not in the former, or vitamin C-synthesizing animals.

The explanation of the dependence of effect on nutritional requirement may be different for the several inhibitor-metabolite pairs for which it has been observed, but at least in the case of pyrithiamine the mechanism has been elucidated (118). Those forms which do not require thiamine, and which are able to thrive in the presence of the drug, possessed a system, presumably enzymic, which split the drug into its pyrimidine and pyridine parts in the same fashion as thiamine may be cleaved at an analogous point into its pyrimidine and thiazole halves. Organisms which could not synthesize thiamine did not possess this ability to destroy the agent, at least not in appreciable amount, and hence were subject to attack.

Another insight into the relationship between nutritional requirement and effectiveness of an agent may be had in the case of organisms which need folic acid. Most such species are rather resistant to the bacteriostasis exerted by sulfonamides. Now, p-aminobenzoic acid is a part of folic acid, and presumably is used to form the vitamin. If the sulfonamides compete with p-aminobenzoic acid in this synthesis, organisms which cannot perform the reaction anyway, and must depend on the medium for their folic acid, would be expected to be less affected by sulfonamides. Although such a correlation between folic acid requirement and sulfonamide resistance actually does exist (167), it is probably only one aspect of susceptibility to these drugs.

(*c*) *Spectrum of antagonistic action of analogs.* Gradations of activity ranging from full vitamin potency to strong antagonism may be observed if several compounds and organisms are examined. This fact must be borne in mind in the formation of any hypothesis for the mode of operation of antagonistic structural analogs. As one examines the data available at present he can see three aspects of this situation. First, a single inhibitory structural analog may either be antagonistic or have weak metabolite activity when tested on the same organism. Which one of the two opposing effects is manifested is determined by the concentration of the agent. For example, Pollack (88) found that α-methyl pantothenic acid would inhibit growth of *L. casei*, but that as the concentration was raised a stimulation of growth similar to that caused by panthothenic acid was produced. Shive and Snell (90) confirmed this finding, and observed similar situations with certain other inhibitory structural analogs of panthothenic acid. The slight vitamin potency of the analogs, weak as it was, was sufficient to prevent the achievement of complete bacteriostasis because

at the high levels of agent which would have been necessary to bring this about, the pantothenic acid activity of the compound made itself felt. Kuhn *et al.* (22) similarly found both inhibitory and stimulatory actions of p-aminobenzamide on bacteria. Likewise, Woolley (128) observed that the p-nitrobenzyl ether of N-acetyl diiodotyrosine not only was an antagonist of thyroxine as measured on tadpoles, but also in the absence of the latter compound, it was weakly active as the hormone. Just as in the previous example, this metabolite activity of the analog at high concentrations limited the realization of total antagonism of the hormone.

Secondly, in a homologous series of compounds, some may show positive metabolite activity and others may be antagonistic. A good illustration of this is found with derivatives of thiamine which differ in the nature of the alkyl side chain at position 2 in the pyrimidine ring. The vitamin has a methyl group at this point, and, of course, has full metabolite potency. The ethyl has less, but still some vitamin activity. However, the butyl compound competes with thiamine for functions of the latter in rats (123). In this example, the shift from metabolite to anti-metabolite activity is more striking than in the usual cases because with changing structure there is a gradual progression from full to partial activity, and then to antagonistic potency.

Thirdly, a single analog may be active in some species as the metabolite, and in others it may be antagonistic. For example, desthiobiotin is able to replace biotin in promoting growth of *Saccharomyces cerevisiae*, but acts in competition with the vitamin in the growth of *Lactobacillus casei* (34, 35, 36). Why this should be so has been explained by Dittmer and du Vigneaud (35) who showed that the yeast converted desthiobiotin to biotin by the insertion of a S atom, while the lactobacillus could not do this. If the former organism had been unable to effect the conversion it might have succumbed to the antagonistic effects of the analog just as did the bacterium. In similar vein, Woolley (118) has shown that when *Endomyces vernalis* was made resistant to pyrithiamine by continued culture in the presence of the drug, it acquired the power to use pyrithiamine in place of thiamine at the same time that it became insusceptible to the inhibitory action of the analog. As was discussed above, resistance was acquired by development of the ability to split pyrithiamine into its pyrimidine and pyridine halves. Since the pyrimidine was the same as that in thiamine, and since for *Endomyces* this fragment possessed full vitamin activity, the organism used it as a growth factor. Similarly, although desoxypyridoxin was a powerful competitor to pyridoxin in the chick (103) it possessed no activity for the rat[1], and for lactic acid bacteria it had weak vitamin potency (168).

The mere ability of some organisms to convert an antagonistic analog into the metabolite as in the examples given above, is not the sole explanation of the varied susceptibility of species. For example, when the S atom of biotin was replaced by an O, the resulting analog, oxybiotin, was found to have vitamin activity equal to that of biotin for some organisms (169), slight potency for others (170), and none at all for still others (171). Indeed, it was actually

[1] Recently some antipyridoxin activity of this compound in rats has been demonstrated. (Porter, C. C, I. Clark and R. H Silber, J. Biol. Chem. 167: 573, 1947).

toxic for certain forms although this toxicity was not demonstrated to be due to competition with biotin. However, no conversion of oxybiotin to biotin could be shown in those microorganisms which used it in place of the naturally occurring vitamin (172). Again, the ethyl analog of pyridoxin (2-ethyl-3-hydroxy-4,5 dimethylol-pyridine) was found to replace the vitamin for tomato roots (173) but to be an anti-pyridoxin in certain fungi (174). It is unlikely that the tomato roots utilized the compound by exchanging the ethyl for a methyl group.

Finally, the ultimate example is provided by the work of Emerson (175) who showed that by long continued cultivation of a *Neurospora* in the presence of sulfanilamide, a strain was secured for which sulfanilamide was a nutritive essential, and p-aminobenzoic acid was an inhibitor of growth. In the parent strain the reverse situation obtained since the sulfonamide was the toxic compound.

(*d*) *Hypotheses concerning mechanism.* Of the various hypotheses proposed to explain why structurally similar compounds should compete with each other in biological systems, the one which fits the observed facts best and which has been most stimulating to further discovery is the one which pictures metabolite and inhibitory analog contending for the specific part of a protein, possibly an enzyme, with which the metabolite normally reacts. Some facts observed experimentally require modifying assumptions in order to conform to this view. One such fact is that subinhibitory concentrations of antagonists such as sulfanilamide or pyrithiamine (117, 115) frequently cause stimulation rather than inhibition so that it is necessary to assume compensatory mechanisms at work in the cell in order to reconcile the observation with the hypothesis. However, when one views the large body of evidence which is best explained by this view of mechanism, and recalls how productive it has been of new findings, he is persuaded that until a better working hypothesis comes along he will hold this one as true. This is especially so when one remembers that rival explanations such as the one that metabolite inhibitors exert a narcotic effect (176) have been rather sterile. Let us then examine the favored hypothesis.

According to this view, a metabolite functions by first combining with a particular cellular unit which is usually regarded as an enzyme or possibly some other specific protein. The union of a substrate with its enzyme would be an analogous reaction, and indeed, the metabolite in this rôle should be regarded as a substrate. The combination then proceeds to pass through a second metabolic reaction or a chain of reactions with a new compound (or substrate) which is transformed and the metabolite is either regenerated or changed into another substance. The cycle of alterations undergone by cozymase in transporting hydrogen would be an example of such a process. Now, in order for the first step to occur, the metabolite must possess certain structural features, and the firmness of combination with the protein is determined in part by these. A similarly constituted compound is able to undergo this combination, but since its structure is not identical with that of the metabolite, its firmness of union will be different. The conjugate of analog and protein is, however, a new and foreign compound

and is unable to proceed through the rest of the cycle of reactions. Like the fabled dog in the manger the analog denies the organism the use of the metabolite and thus creates a deficiency of the latter.

The union of metabolite and protein or of analog and protein is a reversible one usually, so that relative concentrations and relative combining power of metabolite and analog determine which one shall enjoy the protein. Since the metabolite usually has the greater combining power, less of it than of analog will saturate the protein. Therefore, as was pointed out before, the inhibition index is generally greater than 1. As the concentration of analog is increased in the presence of a constant amount of metabolite, a point is reached at which the protein unites with the former rather than with the latter substance, and inhibition of the metabolic function results. If the amount of metabolite is then raised, the analog is dislodged and normal function is resumed. If an analog should be so constituted that it would combine irreversibly with the protein, an inhibitor would result which would call forth signs of deficiency in the organism but these manifestations could not be reversed by increased doses of the metabolite.

What are the concrete pieces of evidence for this hypothesis? The major ones are the following. In the first place, it is fairly well established by numerous studies in enzymology that substrate and enzyme actually do combine reversibly as the first stage in the reaction. Here then is one type of system in which the first reaction in our hypothesis does occur. Indeed, such combinations are known which involve classical metabolites, e.g., the attachment of thiamine pyrophosphate to carboxylase. Furthermore, the competition of structurally similar substrates for a single enzyme has been known for decades, and can be seen to good advantage in the contending of succinate and malonate for succinic dehydrogenase (111). Finally, in non-enzymic reactions between specific proteins and metabolites, solid proof exists that the metabolite can be displaced from the combination by structurally analogous compounds. The best known case is that of the displacement of oxygen in oxyhemoglobin by carbon monoxide. If there are 210 molecules of oxygen for each one of carbon monoxide, the latter is excluded from half the protein and an equal mixture of oxyhemoglobin and carbon monoxyhemoglobin exists. As the relative concentration of carbon monoxide is increased, oxygen is pushed from its combination and carbon monoxide takes its place. As the relative concentration of oxygen increases, the reverse process occurs. Again, biotin combines in molecular proportions with a specific protein, antibiotin or avidin (117, 178). When biotin sulfone is admitted to this system containing the combination, biotin is displaced and appears free in the solution (35). Biotin sulfone not only competes with biotin in this simple system but also functions antagonistically with it in the growth of *Saccharomyces cereviscae*. Another example is the appearance of acetyl choline in solution when sufficient physostigmine is added to a preparation containing cholinesterase and the hormone (179). Thus there is a considerable body of evidence to show that metabolites can be displaced from combination with specific proteins by adequate amounts of certain structural analogs.

Whether they all can be, or whether in fact they all act by combination with proteins or other cell structures, cannot be said at present.

Although many authors have made the tacit assumption that analogs compete with metabolites acting as co-enzymes, there is no direct evidence that they do. In fact, most available data show that the competition is expressed when the metabolite acts as a substrate and not as a co-enzyme. In behaving as a substrate, the metabolite is either degraded or is synthesized into a more complex molecule. This latter compound may then proceed to function as a co-enzyme. For example, pyrithiamine does not compete with thiamine pyrophosphate in the rôle of co-carboxylase. Rather it is probable that the competition between thiamine and pyrithiamine occurs in the synthesis of the co-enzyme from the vitamin.

Intensive study of the mode of action of several inhibitory structural analogs has indicated that it is usually a synthetic or anabolic reaction rather than a degradative or catabolic one which is affected. However, sufficient numbers of cases have not been examined intensively enough to allow one to conclude that this is a general situation. Indeed, the competition between succinate and malonate, or that between thiamine and o-aminobenzyl methyl thiazolium chloride (124) are examples of the participation of inhibitory analogs in degradative reactions. Nevertheless, some insight has been gained into synthetical processes by studies of antagonism, and therefore this aspect of the mechanism of action deserves our attention.

Pantoyl taurine or thiopanic acid is a bacteriostatic analog of pantothenic acid. McIlwain (82) has shown that this agent competes with the vitamin in resting streptococci in an enzymic process concerned with the conjugation of pantothenic acid into a larger molecule in the cells. Although the nature of this conjugate is unknown, good evidence exists for its occurrence in bacteria as well as in animals (180). Therefore, pantoyl taurine competes with pantothenic acid as a substrate used in the synthesis of a larger, and presumably metabolically important, substance.

Again, Lampen and Jones (167) have produced strong evidence that the sulfonamide drugs compete with p-aminobenzoic acid as substrate for the synthesis of pteroyl glutamic acid, a metabolically important conjugate of p-aminobenzoic acid. As was discussed in a previous section, these investigators found that those bacteria which require pteroyl glutamic acid as a nutritive essential, and hence cannot synthesize it, are not susceptible to bacteriostasis by concentrations of sulfonamides which affect other bacteria. Presumably, if the organism does not possess a means of synthesizing PGA, it cannot be harmed by an agent which acts to inhibit this synthesis. Furthermore, PGA antagonises the bacteriostatic action of sulfonamides on certain susceptible species. In other words, when the synthesis of PGA is inhibited, the harm to the organism can be circumvented by supplying in the medium the product of the reaction, viz., PGA. They showed as further evidence that while the antagonism between p-aminobenzoic acid and the sulfonamide is of a competitive nature over a wide range of concentration, the same amount of PGA is required to reverse a large dose of the

sulfonamide as is needed for a small one, i.e., that the antagonism is not competitive. (See Table 3.) Prior to the discovery of p-aminobenzoic acid as a part of PGA, Miller (181) had shown that when bacteria were grown in the presence of sulfonamides, the amount of PGA which they formed was markedly reduced. In the light of all this evidence, the most plausible explanation for the bacteriostasis caused by sulfonamides is that they compete with p-aminobenzoic acid for the synthesis of pteroyl glutamic acid (folic acid). Since some animals need the latter, they are like the PGA-requiring bacteria in that they do not possess the metabolic process with which the sulfonamides interfere.

This hypothesis about mechanism, which may be called the displacement hypothesis, has its roots extended far back into biochemistry and pharmacology. In various forms it has been reiterated since the days of Ehrlich. The competition of structurally similar substrates for specific enzymes reaches back

TABLE 3

*Data of Lampen and Jones showing the growth and antisulfonamide activity of pteroylglutamic acid and p-aminobenzoic acid**

COMPOUND	S. FAECALIS (RALSTON) Half maximum antagonism of sulfadiazine (SD) 1 γ SD per ml.	10 γ SD per ml	100 γ SD per ml	1000 γ SD per ml.	50 PER CENT INCREASE IN TURBIDITY	S FAECALIS R, HALF MAXIMUM GROWTH
p-Aminobenzoic acid	0 003	0 03	0.3	3.0	0 001	Inactive
p-Aminobenzoylglutamic acid	1.	30	300	>300	0.05	Inactive
Pteroylglutamic acid	0.0003	0.0003		0.0003	0.0001	0.00015
Pteroyltriglutamic acid.	0.004	0.004	0.004	0.008	0.008	0 003
Thymine	0.06	0.25	0.25	0.25	0 25	0 3–0.5

* All values are given in micrograms per ml , after 24 hours' incubation.

through the work of Quastel and Wooldridge with succinic dehydrogenase in 1927, at least to 1910 when the inhibition of amylase by various carbohydrates was studied by Wohl and Glimm (182). Many of the current ideas especially as they apply to pharmacology were clearly stated by Clarke in 1937 (183). Woods (20) and Fildes (184), and the contemporary reviewers and investigators have discussed and added to it. It is a concept which has grown with the acquisition of biochemical knowledge.

(*e*) *Competition between structurally dissimilar compounds.* The action of certain agents is antagonized not only by the metabolites to which they bear analogy, but also by structurally unrelated substances. If the subject of mere antagonism of biochemical reactions was to be discussed fully it would take us far afield from the purpose of this review which is to describe the antagonism between structurally similar materials. As others have pointed out (183, 17) two compounds may be antagonistic because they react with each other to form precipitates or biologically inert substances, or because one injures the living

cell in such a way that the other is able to replace the lost function. Now surely the antagonism between the sulfonamides and pteroyl glutamic acid which was discussed above is of this latter variety, and yet it impinges so directly on the phenomenon of competitive antagonism between structural analogs that it must be touched upon here. By the same token we must consider a few more examples but leave to others the treatment of such agents as thiouracil which antagonizes thyroxine by inhibiting its formation. Once again the problem of where to set the limits of a biological phenomenon is plagued by the continuity and interrelationships of nature.

Not only can the growth-inhibiting properties of sulfanilamide be overcome with p-aminobenzoic acid, but they may also be counteracted with methionine or with adenine (60, 185, 186, 205). However, while the action of increasing concentrations of the sulfonamide is negated by proportional increases of p-aminobenzoic acid so that the inhibition index remains constant, this is not true of antagonism between sulfanilamide and either methionine or adenine. With these latter substances as the concentration of drug is increased a point is soon reached beyond which no amount of material will effect reversal. Furthermore, the concentration of adenine or methionine needed to counteract a given quantity of sulfanilamide is far greater than that of p-aminobenzoic acid. Consideration of these facts led Harris and Kohn (60) to postulate that methionine or adenine might be products of a metabolic reaction for which p-aminobenzoic acid was required. This same idea was adopted by Shive and Roberts (186) who elaborated it with a mathematical treatment. Para-aminobenzoic acid was pictured as participating in several biochemical transformations, the product of one being adenine and of another methionine. Some supporting evidence for this postulate was cited, but hypothesis is still ahead of proof.

Other cases of antagonism between structurally dissimilar compounds are known. Thus the microbiostatic effects of phenyl pantothenone are prevented by glutamic acid, proline, or histidine as well as by the structurally related pantothenic acid (95), and the pharmacological manifestations of 3-acetylpyridine in mice are reversed by tryptophane as well as by the analogous nicotinic acid (187). The general features of the antagonism between the unrelated compounds in these cases are similar to those discussed for the sulfonamides and methionine and adenine. Curiously enough, the heterologous reversing agents in most instances are amino acids (3).

For many inhibitory structural analogs there are species in which the related metabolite will not compete with the agent (see section (*a*) above). However, in such organisms the structurally dissimilar reversing agent frequently will counteract the manifestations. For example, although pantothenic acid was unable to reverse the inhibition of growth of *Saccharomyces cerevisae* produced by phenyl pantothenone, glutamic acid, proline, or histidine would do it. Again, while p-aminobenzoic acid was unable to counteract p-aminomethylbenzene sulfonamide (marfanil), methionine was capable of reversing it. From these and similar examples it would seem that the heterologous naturally occurring

reversing agents (metabolites) may be more universally effective than the analogous ones.

Competition between Structurally Analogous Drugs. It is not necessary that one of a pair of antagonistic structural analogs be a metabolite in order that biological competition may be manifested. Two substances entirely foreign to an organism may be shown to contend within it in order to exert a pharmacological action. Thus, some of the effects of morphine in dogs and mice were counteracted by the structurally related allyl normorphine (188). It is relatively certain that neither of these substances occurred in these animals. In view of this situation, the competition in yeast between 2,3-dichloronaphthoquinone and vitamin K may be regarded as the contention of two foreign substances, since vitamin K has not been detected in this organism (151). However, on the other hand, if vitamin K were an important substance to the yeast cell and yet occurred as only a few molecules per cell, its presence could not be detected by present methods. Whether or not structurally similar drugs compete with each other, because they both are analogs of some normal metabolite cannot be decided from existing knowledge.

Applications to Pharmacology. Because the phenomenon of competition between analogous compounds offered a plausible explanation of the mode of action of the sulfonamide drugs, the expectation arose that new agents for combatting infectious diseases might be formed by suitably altering the structures of other metabolites (184, 189, 160). Since initial attempts to do this met with scant success, several students of chemotherapy of infectious diseases were led to doubt the feasibility of such an approach. Nevertheless progress is being made so that there is still reason to cherish the original ideas. For example, 2,3-dichloronaphthoquinone, a useful agent in the chemotherapy of fungal infestations of plant materials (150), was developed empirically, and then shown to compete with vitamin K (151). The sequence of events in this case paralleled those with sulfanilamide and p-aminobenzoic acid. In similar fashion, Kirkwood and Phillips (54) have indicated that the insecticide hexachlorocyclohexane, owes part of its pharmacological action to the fact that it is an analog of inositol, because the vitamin will reverse the yeast-growth-inhibiting powers of the drug. Here again the usefulness of the agent was established before its relationship to a metabolite was pointed out.

Let us now see whether the converse is true; namely, that active drugs can be produced by intentionally modifying some metabolite. McIlwain and Hawking (73) demonstrated that pantoyl taurine, or thiopanic acid, an inhibitory analog of pantothenic acid, could protect rats from a fatal infection by streptococci, but this agent was too weakly active to make it of any value as a practical drug, and the same proved to be true of several derivatives of it. Finally, however, White *et al.* (93, 17) found that the dibromoanilide of thiopanic acid was a rather potent chemotherapeutic agent in animals. Furthermore, phenyl pantothenone, another type of pantothenic acid analog, has proven to be as effective as quinine in the treatment of blood-induced malaria of birds, and likewise to be of value in similarly initiated infestations of man (3, 96). Trial

of phenyl pantothenone in malaria was prompted by Trager's observation (190) that pantothenic acid increased the survival time of the parasites *in vitro*, and hence was probably of importance in their economy. These small advances therefore may serve as straws in the wind to point the direction of future studies. Empiricism will undoubtedly lead to many discoveries in this field, just as it has always done, but a rational approach is commanding some attention.

In pursuing this rational approach, one guidepost must be the selectivity of action of the agents. If an inhibitor is very effective in causing signs of deficiency of the host, its usefulness in combatting a parasite will be small even though it may be quite potent as a growth-suppressant for the latter. Now, several inhibitory structural analogs of metabolites are known which act against certain classes of organisms and not against others. For example, as we have seen earlier, pyrithiamine affects only those species which have a nutritional requirement for thiamine. The sulfonamides affect many microorganisms but do not produce signs of deficiency of p-aminobenzoic acid when given in therapeutic amounts to animals, and there is some evidence that this selectivity is related to the ability to synthesize pteroyl glutamic acid (see above). Phenyl pantothenone, as well as other known antagonists of pantothenic acid, does not cause signs of deprivation of the vitamin in animals, although it is quite active against microorganisms. Acetylpyridine competes with nicotinic acid in animals, but not in bacteria or yeasts (68), and 2,3-dichloronaphthoquinone is far more toxic to fungi and yeasts than to bacteria and animals (151). Although some insight has been gained into the reasons for such discrimination, very little is known about the matter, and if more were understood, therapeutic agents against infectious diseases might be sought more intelligently.

Woolley (189, 159) has underscored the fact that chemotherapy of infectious diseases is only a small part of the chemotherapy of disease, and has shown how inhibitory structural analogs of metabolites may be applied in attempts to treat non-infectious maladies. It must be emphasized that no practically useful drugs have been turned up as a result of these studies, but some experimental models have been set up to explore the possibilities. Thus, it was predicted that an inhibitory analog of tocopherol would cause resorptive interruption of pregnancy in mice, and it was shown that α-tocopherol quinone would bring about such a result (127). By knowing the pharmacological manifestations of a deficiency of a metabolite, the effects of a metabolite analog may be in part anticipated. Again, certain hormones cause disease because they are either produced at an excessive rate or are destroyed abnormally slowly. It might be possible to neutralize the effects of the excess hormone by administration of an inhibitory analog. Such analogs of thyroxine have been formed, and shown to counteract toxic doses of the hormone given to tadpoles (128). Furthermore, abundant experimentation is being carried out to learn whether some of the agents antagonistic to histamine and effective against anaphylactic shock and against certain allergic reactions, may be viewed as competitive structural analogs of histamine (17).

The idea that various drugs act by combining with specific portions of the cell, and thus preventing the normal metabolites from reacting, goes back to the time

of Ehrlich. Since the evidence for this view, and the examples of it have been lucidly expounded by Clark (183), and recent developments especially in relation to participation of sulfhydryl compounds in these processes, by Roblin (17), no discussion will be included here.

Applications to Biochemistry. Inhibitory structural analogs of metabolites have appealed to many investigators as powerful tools with which to examine biochemical reactions. By specifically interfering with one process it may be possible to discover secrets of metabolism which have eluded us. As with all new tools, however, there must be a period of learning how to use them, and so with these there has been some fumbling. As a result much of the earlier work has been richer in speculation and hypothesis than in pertinent experimental observation. Let us scan the work along this line then more with the idea of seeing the type of use to which analogs have been and are being put, rather than with the hope of reviewing each individual contribution.

Several authors have concluded from their data on antagonisms that a given metabolite has a variety of functions. Now this view is not new, but in the case of some vitamins, it added to our knowledge. Thus, Shive and Roberts (186) have postulated several functions for p-aminobenzoic acid because of the existence and behaviour of several antagonists to sulfonamide drugs. Previously, Woolley (68) had considered the possibility of varied rôles for nicotinic acid, because of the way in which acetylpyridine affected animals and microorganisms.

Whatever may be the merits of these postulations there is no doubt that different types of inhibitory analogs of the same metabolite exert qualitatively different biological effects. Thus although thiopanic acid (pantoyl taurine) and phenyl pantothenone are two differing but close analogs of pantothenic acid, and both compete with the vitamin in the growth of bacteria, their qualitative effects are rather dissimilar as witnessed by the fact that thiopanic acid is ineffective in the treatment of malaria, or in the prevention of growth of non-pantothenic acid-requiring microorganisms while phenyl pantothenone possesses these powers. Or again, there are four types of structural analogs of vitamin K, represented by 3,3′-methylenebis-(4-hydroxycoumarin), iodinin (a phenazine oxide), α-tocopherol quinone, and 2,3-dichloronaphthoquinone. The action of each of these compounds may be reversed by vitamin K, but of them only the first influences the prothrombin level of plasma, and whereas the first causes generalized hemorrhages in animals, the third induces bleeding only from the reproductive tract, and that only in pregnant animals (127). The first is only weakly active in the inhibition of growth of bacteria, and in this process is not antagonized by vitamin K, while the fourth is highly potent against certain microorganisms in which its effect may by antagonized by vitamin K.

Obscure interrelationships between metabolites have been indicated by observations with their analogs. Here too as in the former cases, other types of evidence frequently pointed to similar conclusions. Thus, α-tocopherol quinone, an analog both of vitamin K and of tocopherol, called forth pharmacological signs similar to those seen in deficiency of the latter vitamin, but these manifestations were negated only by the former (127). Some metabolic relationship

between tocopherol and vitamin K seemed indicated. The experiments with sulfonamides led to the postulation of p-aminobenzoic acid as a metabolite (20) and this hypothesis was amply verified subsequently. The fact that tryptophane intake reduced the requirement for nicotinic acid, which has led to the postulation of tryptophane as the precursor of nicotinic acid (191), was discovered without the conscious use of analogs, but the interrelationship was exemplified again by the use of 3-acetylpyridine (187). Furthermore, the original demonstration may have depended on the existence of an analog of the vitamin in the corn used in the diet (192). Shive and Macow (12) have felt that the effects of pantothenic acid on the antagonism between aspartic acid and its hydroxy derivative indicated that aspartic acid was the precursor of the vitamin.

Because of the greater activity of cocarboxylase as compared to thiamine in antagonizing the microbiostatic effects of pyrithiamine, Sarett and Cheldelin (120) concluded that the vitamin was probably attached to the protein part of thiamine-containing enzymes before the pyrophosphate ester was formed. Similar findings with analogs of riboflavin led Sarett (109) to the same view with respect to the functioning of that vitamin, namely, that the metabolite was joined to the protein before phosphorylation converted the former into a coenzyme. However, these would not constitute unique explanations of the experimental data.

Several studies have been made of the use of inhibitory structural analogs of metabolites for the specific inhibition of various enzymes. Not only have new inhibitors of old enzymes been found, but also the existence of previously unrecognized enzymic reactions has been discovered. Sealock and Goodland (124) showed that the enzymic splitting of thiamine by carp thiaminase (193, 194) could be inhibited competitively by the thiamine analog o-aminobenzyl-methylthiazolium chloride, and that the oxidation of tyrosine by liver slices, a process catalyzed by ascorbic acid, was inhibited by glucoascorbic acid (15). Likewise von Euler and Ahlström (195) reported an interfering action of pyridine-3-sulfonic acid in the dehydrogenation of lactate by cozymase containing systems. The interpretation of these findings was made more complex by the fact that several unrelated compounds (e.g., salicylic acid) could replace pyridine-3-sulfonic acid. Perhaps some of the obscurity in interpretation of these results resided in the fact that the analog was being used to compete with a coenzyme rather than for a substrate. An investigation of the effects of nicotinic acid analogs on the enzymic synthesis of cozymase might provide more interpretable data. (Compare the discussion of mechanism above). Paget and Vittu (196) have reported that phospho-mono-esterase of erythrocytes is retarded in action by sulfanilamide and that this effect is reversed by p-aminobenzoic acid. In view of this finding a search for the function of the latter metabolite in the enzyme-system might prove profitable.

However, the mere inhibition of an enzymic reaction by a metabolite analog is not trustworthy evidence that the related metabolite is actually an integral part of the system. For example, although sulfanilamide interferes with the action of carbonic anhydrase (197), there are no data to indicate that p-amino-

benzoic acid is concerned with the enzymic process. The effect of the analog is not reversed by p-aminobenzoic acid, and furthermore, other sulfonamides which compete with the metabolite do not inhibit the action of the enzyme. In this instance sulfanilamide seems to possess a biological property not related to its competition with the structurally similar p-aminobenzoic acid. Many chemical compounds are harmful to enzymic reactions, so that it is understandable that metabolite analogs may have effects other than those related to their competitive behaviour.

A new enzymic process, namely, that of the conversion of pantothenic acid into a conjugated form in streptococcal cells was discovered by McIlwain through the use of the analog thiopanic acid (83). This reaction of the vitamin was intimately coupled with glycolysis but by judicious use of inhibitors, the pantothenic acid conjugating system could be differentiated from the glucolytic one. In similar fashion, Woolley (118) recognized a pyrithiamine-splitting system in certain microorganisms, the occurrence of which was correlated with the ability to synthesize thiamine.

The observation that a pair of structurally similar metabolites, which both occur in the same individual, may be competitively antagonistic to each other has led to the hypothesis that certain metabolic functions are regulated through the intermediation of this antagonism. For example, Granick and Gilder (101) considered that metabolism of porphyrins was in part regulated by the competition between protoporphyrin and copro-porphyrin. These two biologically important compounds were known to exist in the same individual, and yet in the growth of *Hemophilus influenzae* they were competitors. A similar view for the estrogens and androgens was mentioned briefly above (115). Possibly the antagonisms between analogous amino acids, such as that between glycine and alanine (7) may be viewed from this angle.

Compounds have been discovered in some species which are harmful to other living things, and which can be rendered harmless by an excess of structurally analogous metabolites. In other words, inhibitory structural analogs of metabolites are formed by one type of organism and act against other types. Thus, iodinin, a hydroxy phenazine oxide, is an antibiotic pigment formed by *Chromobacterium iodinum*. It competes with the structurally related vitamin K in the prevention of growth of streptococci (149). Again, a substance has been found in corn (maize) which is able to cause signs of pellagra in mice (192). The action of the compound was reversed either by nicotinic acid or by tryptophane. Because of these facts, and because of the chemical properties of the active substance, it was postulated that the agent might be an analog of nicotinic acid.[1]

[1] Since this paper was prepared the pellagragenic agent has been identified as indole-acetic acid (200). This compound may be viewed as derived from nicotinic acid by conjugation with a benzene ring and extrusion of a carbon atom from the pyridine ring. It is likewise an even better analog of tryptophane. Just as in the case of α-tocopherol quinone, which is related to both vitamin E and vitamin K, indole-acetic acid is at the same time an analog of two metabolites. In the latter instance either metabolite will reverse the effects of the analog.

This last example lacks force because the nature of the agent is still conjectural, but a better defined case of production of disease by a metabolite analog exists. *Fusarium lycopersici*, the cause of a common disease of tomatoes, produces a toxic substance which will reproduce the signs of the disease in plants. This agent has been isolated in pure form (198) and shown to be a tripeptide isomeric with serylglycylaspartic acid (199). The tomato-wilting powers of this peptide are negated by the related serylglycylglutamic acid, or glutathione, or by strepogenin, and these latter peptides have been shown to function as required nutrients for several species of bacteria (199). The Fusarium probably is able to cause disease by forming an inhibitory structural analog of a metabolite. The logical step would be to treat the infection with the metabolite. The elucidation of the structure of the antagonist in this case was the key for solving the probable nature of the unknown metabolite, strepogenin.

Some Relationships of Structure to Antagonistic Activity. Generalizations have been made about the types of structural alteration which will convert metabolites into antagonistic agents (189). Despite some adverse criticisms (e.g. 18), and in the face of acknowledged shortcomings and limitations (189, 159) the ideas contained in these generalizations have continued to lead to the production of new metabolite-inhibitors. The basis for these rules may be seen from table 2. This is that alteration of the structure of a metabolite in certain defined ways results in the formation of agents which compete with it. Some of these types of alteration are: the replacement of a carboxyl group with a sulfonic acid or sulfonamide radical, or by a ketone; the substitution of one atom for another in the ring system of the metabolite; and the replacement of alkyl side chains of aromatic ring systems with halogens. These are only a beginning in the classification of generalizations about the relationships of chemical constitution to antagonistic activity, because many examples could be cited of other types of alterations in structure which lead to inhibitory analogs (3). Furthermore, Woolley (189) has pointed out that there is no single way in which the structure of a vitamin or hormone must be altered in order to obtain a competitor to it. As we saw in the previous section, there may be pronounced qualitative differences in effects produced by different types of analogs of the same metabolite. As knowledge about the phenomenon of antagonism between structurally similar compounds accumulates it is probable that these postulations will be both modified and expanded.

REFERENCES

(1) Goodman, L., A. Gilman and N. Hart. Fed. Proc. 2: 80, 1943.
(2) Woolley, D. W. J. Biol. Chem. 152: 225, 1944.
(3) Woolley, D. W. The Harvey Lectures, Series *XLI*, 1945–46.
(4) Roblin, R. O., Jr., J. O. Lampen, J. P. English, Q. P. Cole and J. R. Vaughan. J. Am. Chem. Soc. 67: 290, 1945.
(5) Nielsen, N. Naturwissenschaften 31: 146, 1943.
(6) Sarett, H. P. and V. H. Cheldelin. J. Bact. 49: 31, 1945.
(7) Snell, E. E. and B. M. Guirard. Proc. Nat. Acad. Sci. 29: 66, 1943.
(8) McIlwain, H. J. Chem. Soc. 75, 1941.
(9) Spizizen, J. J. Infect. Dis. 73: 212, 1943.

(10) GLADSTONE, G. P. Brit J Exper. Path 20: 189, 1939.
(11) HUTCHINGS, B. L. AND W. H PETERSON. Proc. Soc Exper. Biol. and Med. 52: 36 1943.
(12) SHIVE, W. AND J. MACOW. J. Biol. Chem. 162: 451, 1946.
(13) WOOLLEY, D. W. AND L. O. KRAMPITZ. J. Exper. Med. 78: 333, 1943.
(14) WOOLLEY, D. W. Fed. Proc. 3: 97, 1944.
(15) LAN, T. H. AND R R. SEALOCK. J. Biol. Chem. 155: 483, 1944; Gibson Island Conferences of Am. Assoc. Advancement of Science, July 1944.
(16) HENRY, R. J. Bact. Rev. 7: 175, 1943.
(17) ROBLIN, R. O., JR. Chem. Rev. 38: 255, 1946.
(18) WELCH, A. D. Physiol. Rev. 25: 687, 1945.
(19) NORTHLY, E. H. Chem. Rev. 27: 85, 1940.
(20) WOODS, D. D. Brit. J. Exper. Path. 21: 74, 1940
(21) HIRSCH, J. Science 96: 139, 1942.
(22) KUHN, R., E. F. MÖLLER, G. WENDT AND H. BEINERT. Ber. 75: 711, 1942.
(23) PETERS, L. J. Pharmacol. 79: 32, 1943.
(24) MÖLLER, E. F. AND K. SCHWARZ. Ber. 74B: 1612, 1941.
(25) SANDGROUND, J. H. AND C. R. HAMILTON. J. Pharmacol. and Exper. Therap. 78: 109, 1943.
(26) SANDGROUND, J. H. AND C. R. HAMILTON. J. Pharmacol. and Exper. Therap. 78: 203, 1943.
(27) SANDGROUND, J. H. J. Pharmacol. and Exper. Therap. 78: 209, 1943.
(28) SANDGROUND, J. H. AND C. R. HAMILTON. J. Lab. and Clinical Med. 28: 1821, 1943.
(29) JOHNSON, O. H., D. E. GREEN AND R. PAULI. J. Biol. Chem. 153: 37, 1944.
(30) WYSS, O., M. RUBIN AND F. B. STRANDSKOV. Proc. Soc. Exper. Biol. and Med. 52: 155, 1943.
(31) MARTIN, A. R. AND F. L. ROSE. Biochem. J. 39: 91, 1945.
(32) ANDREWES, C. H., H. KING, M. VAN DEN ENDE AND J. WALKER. Lancet 1: 777, 1944.
(33) AUHAGEN, E. Ztschr. physiol. Chem. 274: 48, 1942.
(34) DITTMER, K., D. B. MELVILLE AND V. DU VIGNEAUD. Science 99: 203, 1944.
(35) DITTMER, K. AND V. DU VIGNEAUD. Science 100: 129, 1944.
(36) LILLY, V. G. AND L. H. LEONIAN. Science 99: 205, 1944.
(37) DU VIGNEAUD, V. Chem. Eng. News 23: 623, 1945.
(38) BROWN, G. B. AND V. DU VIGNEAUD. J. Biol. Chem. 163: 761, 1946.
(39) ENGLISH, J. P., R. C. CLAPP, Q. P. COLE, I. H. HALVERSTADT, J. O. LAMPEN AND R. O. ROBLIN, JR. J. Am. Chem. Soc. 67: 295, 1945.
(40) AXELROD, A. E., J. DEWOODY AND K. HOFMANN. J. Biol. Chem. 163: 771, 1946.
(41) KESTON, A. S. AND S. B. WORTIS. Proc. Soc. Exper. Biol. and Med. 61: 439, 1946.
(42) KUHN, R. Angew. Chem. 53: 1, 1940.
(43) BOREK, E., P. SHEINESS AND H. WAELSCH. Fed. Proc. 5: 123, 1946.
(44) BOREK, E., H. K. MILLER, P. SHEINESS AND H. WAELSCH. J. Biol. Chem. 163: 347, 1946.
(45) MORRIS, H. C. AND C. A. DRAGSTEDT. Proc. Soc. Exper. Biol. and Med. 59: 31, 1945.
(46) LANDAU, S. W. AND L. N. GAY. Bull. Johns Hopkins Hosp. 74: 55, 1944.
(47) JADASSOHN, W., H. E. FIERZ-DAVID AND H. VOLLENWEIDER. Schweiz. Med. Wchnschr. 73: 122, 1943.
(48) JADASSOHN, W., H. E. FIERZ-DAVID AND H. VOLLENWEIDER. Helv. Chim. Acta 27: 1384, 1944.
(49) EMMELIN, N., G. S. KAHLSON AND K. LINDSTRÖM. Acta Physiol. Scand. 3: 39, 1941.
(50) YOUNG, R. H. AND R. P. GILBERT. J. Allergy 12: 235, 1941.
(51) GILBERT, A. J. AND F. GOLDMAN. Proc. Soc. Exper. Biol. and Med. 44: 458, 1940.
(52) CHAUCHARD, B. AND P. CHAUCHARD. Compt. rend. soc. biol. 137: 708, 1943.
(53) SKOOG, F., C. L. SCHNEIDER AND P. MALAN. Am. J. Botany 29: 568, 1942.
(54) KIRKWOOD, S. AND P. H. PHILLIPS. J. Biol. Chem. 163: 251, 1946.

(55) Doudoroff, M. Proc. Soc. Exper. Biol. and Med. **53:** 73, 1943.
(56) Bonner, D., E. L. Tatum and G. W. Beadle. Arch. Biochem. **3:** 71, 1943.
(57) Fox, S. W., M. Fling and G. N. Bollenbach. J. Biol. Chem. **155:** 465, 1944.
(58) Doermann, A. H. Arch. Biochem. **5:** 373, 1944.
(59) Dyer, H. M. J. Biol. Chem. **124:** 519, 1938.
(60) Harris, J. S. and H. I. Kohn. J. Pharmacol. **73:** 383, 1941.
(61) McIlwain, H. Brit. J. Exper. Path. **21:** 136, 1940.
(62) Matti, J., F. Nitti, M. Morel and A. Lwoff. Ann. inst. Pasteur **67:** 240, 1941.
(63) Möller, E. F. and L. Birkofer. Ber. **75:** 1108, 1118, 1942.
(64) McIlwain, H. Nature **146:** 653, 1940.
(65) Erlenmeyer, H. and W. Würgler. Helv. Chim. Acta **25:** 249, 1942.
(66) Woolley, D. W., F. M. Strong, R. J. Madden and C. A. Elvehjem. J. Biol. Chem. **124:** 715, 1938.
(67) Woolley, D. W. and A. G. C. White. Proc. Soc. Exper. Biol. and Med. **52:** 106, 1943.
(68) Woolley, D. W. J. Biol. Chem. **157:** 455, 1945.
(69) Erlenmeyer, H., H. Bloch and H. Kiefer. Helv. Chim. Acta **25:** 1066, 1942.
(70) Snell, E. E. J. Biol. Chem. **139:** 975, 1941.
(71) Kuhn, R., T. Wieland and E. F. Möller. Ber. **74:** 1605, 1941.
(72) McIlwain, H. Biochem. J. **36:** 417, 1942.
(73) McIlwain, H. and F. Hawking. Lancet **1:** 459, 1943.
(74) Snell, E., L. Chan, L. Spiridanoff, E. L. Way and C. D. Leake. Science **97:** 168, 1943.
(75) Unna, K. Proc. Soc. Exper. Biol. and Med. **54:** 55, 1943.
(76) Barnett, J. W. and F. A. Robinson. Biochem. J. **36:** 357, 364, 1942.
(77) McIlwain, H. Brit. J. Exper. Path. **23:** 95, 1942.
(78) Rosenthal, Von H. and C. A. Grob. Ztschr. Vitaminforsch. **17:** 27, 1946.
(79) McIlwain, H. Brit. J. Exper. Path. **24:** 203, 1943.
(80) McIlwain, H. Brit. J. Exper. Path. **24:** 212, 1943.
(81) McIlwain, H. Biochem. J. **38:** 97, 1944.
(82) McIlwain, H. and D. E. Hughes. Biochem. J. **38:** 187, 1944.
(83) McIlwain, H. and D. E. Hughes. Biochem. J. **39:** 133, 1945.
(84) Snell, E. E. and W. Shive. J. Biol. Chem. **158:** 551, 1945.
(85) McIlwain, H. Biochem. J. **39:** 329, 1945.
(86) McIlwain, H. Biochem. J. **39:** 279, 1945.
(87) Pfaltz, H. Ztschr. Vitaminforsch. **13:** 236, 1943.
(88) Pollack, M. A. J. Am. Chem. Soc. **65:** 1335, 1943.
(89) Shive, W. and E. E. Snell. Science **102:** 401, 1945.
(90) Shive, W. and E. E. Snell. J. Biol. Chem. **160:** 287, 1945.
(91) Madinaveitia, J., A. R. Martin, F. L. Rose and G. Swain. Biochem. J. **39:** 85, 1945.
(92) Winterbottom, R., J. W. Clapp, W. H. Miller, J. P. English and R. O. Roblin, Jr. To be published (compare reference 17).
(93) White, H. J., M. Lee, E. R. Jackson, A. Himes and C. M. Alverson. To be published (compare reference 17).
(94) Woolley, D. W. and M. L. Collyer. J. Biol. Chem. **159:** 263, 1945.
(95) Woolley, D. W. J. Biol. Chem. **163:** 481, 1946.
(96) Marshall, E. K. Fed. Proc. **5:** 298, 1946.
(97) Martin, G. J., L. J. Lewis and H. Urist. Abstracts of Papers, 109th Meeting. Am. Chem. Soc. 21B, 1946.
(98) Beerstecher, E. and W. Shive. J. Biol. Chem. **164:** 53, 1946.
(99) du Vigneaud, V., H. McKennis, Jr., S. Simmonds, K. Dittmer and G. B. Brown. J. Biol. Chem. **159:** 385, 1945.
(100) Dittmer, K., G. Ellis, H. McKennis, Jr. and V. du Vigneaud. J. Biol. Chem. **164:** 761, 1946.

(101) Granick, S. and H. Gilder. Science 101: 540, 1945.
(102) Keresztesy, J. C., D. Laszlo and C. Leuchtenberger. Cancer Research 6: 128, 1946.
(103) Ott, W. H. Proc. Soc. Exper. Biol. and Med. 61: 125, 1946.
(104) Kuhn, R., F. Weygand and E. F. Möller. Ber. 76: 1044, 1943.
(105) Emerson, G. A. and M. Tishler. Proc. Soc. Exper. Biol. and Med. 55: 184, 1944.
(106) Woolley, D. W. J. Biol. Chem. 154: 31, 1944.
(107) Emerson, G. A., E. Wurtz and O. H. Johnson. J. Biol. Chem. 160: 165, 1945.
(108) Sarett, H. P. Fed. Proc. 4: 101, 1945.
(109) Sarett, H. P. J. Biol. Chem. 162: 87, 1946.
(110) Foster, J. W. J. Bact. 48: 97, 1944.
(111) Quastel, J. H. and W. R. Wooldridge. Biochem. J. 21: 1224, 1927.
(112) Everett, G. and J. C. Krantz, Jr. Proc. Soc. Exper. Biol. and Med. 55: 220, 1944.
(113) Korenchevsky, V. Arch. intern. pharmacodynamie 70: 411, 1945.
(114) Armstrong, W. D., M. Knowlton and M. Gouze. Endocrinology 36: 313, 1945.
(115) Woolley, D. W. Advances in Enzymology 6: 129, 1946.
(116) Woolley, D. W. and A. G. C. White. J. Biol. Chem. 149: 285, 1943.
(117) Woolley, D. W. and A. G. C. White. J. Exper. Med. 78: 489, 1943.
(118) Woolley, D. W. Proc. Soc. Exper. Biol. and Med. 55: 179, 1944.
(119) Wyss, O. J. Bact. 46: 483, 1943.
(120) Sarett, H. P. and V. H. Cheldelin. J. Biol. Chem. 156: 91, 1944.
(121) Robbins, W. J. Proc. Natl. Acad. Sci. U. S. 27: 419, 1941.
(122) Soodak, M. and L. R. Cerecedo. J. Am. Chem. Soc. 66: 1988, 1944.
(123) Emerson, G. A. and P. Southwick. J. Biol. Chem. 160: 169, 1945.
(124) Sealock, R. R. and R. L. Goodland. J. Am. Chem. Soc. 66: 507, 1944.
(125) Buchman, E. R., E. Heegaard and J. Bonner. Proc. Natl. Acad. Sci. U. S. 26: 561, 1940.
(126) Hitchings, G. H., E. A. Falco and M. B. Sherwood. Science 102: 251, 1945.
(127) Woolley, D. W. J. Biol. Chem. 159: 59, 1945.
(128) Woolley, D. W. J. Biol. Chem. 164: 11, 1946.
(129) Fildes, P. Biochem. J. 32: 1600, 1938.
(130) Bloch, H. and H. Erlenmeyer. Helv. Chim. Acta 25: 694, 1062, 1942.
(131) Anderson, T. F. Science 101: 565, 1945.
(132) Woods, D. D. Biochem. J. 36: 4, 1942.
(133) Strandskov, F. B. and O. Wyss. J. Bact. 50: 237, 1945.
(134) Stahmann, M. A., C. F. Huebner and K. P. Link. J. Biol. Chem. 138: 513, 1941.
(135) Link, K. P. Harvey Lectures Series XXXIX, 1944.
(136) Quick, A. J. Physiol. Rev. 24: 297, 1944.
(137) Davidson, C. S. and H. MacDonald. New England J. Med. 229: 353, 1943.
(138) Lucas, S. P. and P. M. Aggeler. Proc. Soc. Exper. Biol. and Med. 56: 36, 1944.
(139) Brodie, D. C., W. A. Hiestand and G. L. Jenkins. J. Am. Pharm. Assoc. 34: 73, 1940.
(140) Cromer, H. E., Jr. and N. W. Barker. Proc. Staff Meetings Mayo Clinic 19: 217, 1944.
(141) Lehmann, J. Lancet 1: 661, 1943.
(142) Overman, R. S., M. A. Stahmann and K. P. Link. J. Biol. Chem. 145: 155, 1942.
(143) Shapiro, S., M. H. Redish and H. A. Campbell. Proc. Soc. Exper. Biol. and Med. 52: 12, 1943.
(144) Meunier, P. and C. Mentzer. Bull. Soc. chim. biol. 25: 80, 1943.
(145) Glavind, J. and K. F. Jansen. Acta Physiol. Scand. 8: 173, 1944.
(146) Overman, R. S., M. A. Stahmann, C. F. Huebner, W. R. Sullivan, L. Spero, D. G. Doherty, M. Ikawa, L. Graf, S. Roseman and K. P. Link. J. Biol. Chem. 153: 5, 1944.
(147) Mentzer, C. and P. Meunier. Bull. Soc. chim. biol. 25: 379, 1943.

(148) Meunier, P., C. Mentzer, Buu-Hoï and P. Cagniant. Bull. Soc. chim. biol. **25**: 384, 1943.
(149) McIlwain, H. Biochem. J. **37**: 265, 1943.
(150) Ter Horst, W. P. and E. L. Felix. Ind. Eng. Chem. **35**: 1255, 1943.
(151) Woolley, D. W. Proc. Soc. Exper. Biol. and Med. **60**: 225, 1945.
(152) Smith, C. C., R. Fradkin and M. D. Lackey. Proc. Soc. Exper. Biol. and Med. **61**: 398, 1946.
(153) Carrara, G. Ricerca sci. **14**: 250, 1943.
(154) Miller, J. K. J. Pharmacol. **71**: 14, 1941.
(155) Kuhn, R., E. F. Möller and G. Wendt. Ber. **76**: 405, 1943.
(156) Drell, W. and M. S. Dunn. J. Am. Chem. Soc. **68**: 1868, 1946.
(157) Wright, L. D. and H. R. Skeggs. Arch. Biochem. **10**: 383, 1946.
(158) Kalnitsky, G. and E. S. G. Barron. (In press.)
(159) Woolley, D. W. in Green, D. E. Currents in Biochemical Research. Interscience Publishers, Inc., New York, 1946.
(160) McIlwain, H. Biol. Rev. Cambridge Phil. Soc. **19**: 135, 1944.
(161) Tracy, A. H. and R. C. Elderfield. J. Org. Chem. **6**: 54, 1941.
(162) Rubbo, S. D. and J. M. Gillespie. Nature **146**: 838, 1940.
(163) Blanchard, K. C. J. Biol. Chem. **140**: 919, 1941.
(164) Angier, R. B., J. H. Boothe, B. L. Hutchings, J. H. Mowat, J. Semb, E. L. R. Stockstad, Y. SubbaRow and C. W. Waller. Science **103**: 667, 1946.
(165) Overman, R. S., J. B. Field, C. A. Baumann and K. P. Link. J. Nutrition **23**: 589, 1942.
(166) Tamura, J. T. J. Bact. **47**: 529, 1944.
(167) Lampen, J. O. and M. J. Jones. J. Biol. Chem. **164**: 485, 1946.
(168) Möller, E. F., O. Zima, F. Jung and T. Moll. Naturwissenschaften **27**: 228, 1939.
(169) Pilgrim, F. J., A. E. Axelrod, T. Winnick and K. Hofmann. Science **102**: 35, 1945.
(170) Hofmann, K., R. H. McCoy, J. R. Felton, A. E. Axelrod and F. J. Pilgrim. Arch. Biochem. **7**: 393, 1945.
(171) Perlman, D. Unpublished data.
(172) Hofmann, K. and T. Winnick. J. Biol. Chem. **160**: 449, 1945.
(173) Robbins, W. J. Am. J. Botany **29**: 241, 1942.
(174) Robbins, W. J. and R. Ma. Bull. Torrey Botan. Club **69**: 342, 1942.
(175) Emerson, S. and J. E. Cushing. Fed. Proc. **5**: 379, 1946.
(176) Johnson, F. H. Science **95**: 104, 1942.
(177) Eakin, R. E., E. E. Snell and R. J. Williams. J. Biol. Chem. **140**: 535, 1941.
(178) Woolley, D. W. and L. G. Longsworth. J. Biol. Chem. **142**: 285, 1942.
(179) Krayer, O., A. Goldstein and F. L. Plachte. J. Pharmacol. **80**: 8, 1944.
(180) Neal, A. L. and F. M. Strong. J. Am. Chem. Soc. **65**: 1659, 1943.
(181) Miller, A. K. Proc. Soc. Exper. Biol. and Med. **57**: 151, 1944.
(182) Wohl, A. and E. Glimm. Biochem. Ztschr. **27**: 349, 1910.
(183) Clark, A. J. in Handbuch der experimentellen Pharmakologie. Vol. IV, Springer, Berlin, 1937.
(184) Fildes, P. Lancet **1**: 955, 1940.
(185) Martin, G. J. and C. V. Fisher. J. Biol. Chem. **144**: 289, 1942.
(186) Shive, W. and E. C. Roberts. J. Biol. Chem. **162**: 463, 1946.
(187) Woolley, D. W. J. Biol. Chem. **162**: 179, 1946.
(188) Unna, K. J. Pharmacol. and Exper. Therap. **79**: 27, 1943.
(189) Woolley, D. W. Science **100**: 579, 1944.
(190) Trager, W. J. Exper. Med. **74**: 441, 1941.
(191) Rosen, F., J. W. Huff and W. A. Perlzweig. J. Biol. Chem. **163**: 343, 1946.
(192) Woolley, D. W. J. Biol. Chem. **163**: 773, 1946.
(193) Woolley, D. W. J. Biol. Chem. **141**: 997, 1941.

(194) Krampitz, L. O. and D. W. Woolley. J. Biol. Chem. **152**: 9, 1944.
(195) Euler, H. v. and L. Ahlström. Ztschr. physiol. Chem. **279**: 176, 1943.
(196) Paget, M. and C. Vittu. Compt. rend. soc. biol. **138**: 1066, 1944.
(197) Mann, T. and D. Keilin. Nature **146**: 164, 1940.
(198) Plattner, P. A. and N. Clauson-Kaas. Helv. Chim. Acta **28**: 188, 1945.
(199) Woolley, D. W. J. Biol. Chem. **166**: 783, 1946.
(200) Kodicek, E., K. J. Carpenter and L. J. Harris. Lancet **2**: 491, 1946.
(201) Tatum, E. L. and T. T. Bell. Am. J. Botany **33**: 15, 1946.
(202) Nielsen, N., V. Hartelius and G. Johansen. Naturwissenschaften **32**: 294, 1944.
(203) Euler, H. v. and P. Karrer. Helv. Chim. Acta **29**: 353, 1946.
(204) Mitchell, H. K. and M. B. Houlahan. Am. J. Botany **33**: 31, 1946.
(205) Schopfer, W. H. and M. Guilloud. Helv. Physiol. Pharmacol. Acta **4**: C24, 1946.

PHYSIOLOGICAL REVIEWS

VOL. 27 JULY, 1947 No. 3

THE BIOLOGICAL SIGNIFICANCE OF HYALURONIC ACID AND HYALURONIDASE

KARL MEYER

The Institute of Ophthalmology of Presbyterian Hospital and the College of Physicians and Surgeons of Columbia University, New York, New York

Hyaluronic acid is a mucopolysaccharide acid which in animal tissues seems to bind water in interstitial spaces. It further holds cells together in a jelly-like matrix and serves as a lubricant and shock-absorber in joints. It is desaggregated and depolymerized by the action of the enzyme hyaluronidase. In this review the effort is being made to discuss aspects most essential to an understanding of the biological rôle of hyaluronic acid and hyaluronidase.

Duran-Reynals in 1942 in a paper entitled "Tissue permeability and the spreading factors in infection" published the first comprehensive review dealing with hyaluronic acid, hyaluronidase and the spreading factors. Current literature has been briefly reviewed in 1942 (15B) and in 1944 (44A).

I. HYALURONIC ACID. A. *Occurrence.* Hyaluronic acid was first isolated from bovine vitreous humor (63). Subsequently it was isolated from human umbilical cord (64), from bovine aqueous humor and from vitreous humor of pig eyes (65), from groups A and C hemolytic streptococci (35), from bovine and human synovial fluid (66), from a number of mesenchymal tumors like fowl leucosis (33), the pleura and peritoneal fluid of a case of mesenthelioma (55), from two cases of human synovioma (unpublished), from Rous and Fuyinami tumors of chicken (71), and from pig skin (56).

Hyaluronic acid probably occurs in the nucleus pulposus of the intervertebral disc (unpublished) and in small concentrations in connective tissue, although it has not been isolated from these sources.

B. *Chemistry.* For the isolation of the acid from various sources some modifications of the same general principles have been described. From fluids like vitreous humor, synovial fluid and some tumor fluids, from which the purest samples have been obtained, the acid is first precipitated as a salt of accompanying proteins. For this purpose the fluids are diluted with 2–5 volumes of cold water and acidified to a pH of about 4 with 50 per cent acetic acid with vigorous stirring. With vitreous humor the precipitation may be incomplete until the protein content is increased by addition of horse or cattle serum. The "mucin clot" is left standing at 0° for 24 hours, is then washed with cold water and extracted with a 5 to 10 per cent solution of sodium, potassium, or calcium acetate, depending upon which salt is to be desired. The extraction may be hastened by adding alkali to a pH of 9.0.

Solid tissues like umbilical cord or some tumors are chilled, ground and desiccated with acetone. The dry material is extracted repeatedly with 5 to 10 per cent of an acetate. The crude acid is acidified to pH 4 and precipitated with 1.5

volume of ethanol. The centrifuged and washed precipitate is again extracted with acetate.

The further purification of this extract and the extracts of the protein salts described above are, from here on, identical. The bulk of the protein is removed by stirring with a chloroform-amylalcohol mixture for one hour followed by separation of the fluid from the gel by centrifugation. The procedure is repeated until a gel is no longer formed. Further nitrogenous impurities are removed either by adsorption on $Zn(OH)_2$ at pH 7.2 (adding 10 per cent zinc acetate and neutralizing) or by adsorption on Lloyd's reagent at a pH of about 4. The supernatant is precipitated with 1.5 volumes of ethanol at pH 4. The stringy mass is again centrifuged, washed with alcohol, redissolved in a minimum amount of water, precipitated by alcohol in the presence of acetate, washed with alcohol and ether and dried in vacuo over P_2O_5. In earlier preparations in this laboratory, the free acid was precipitated by pouring the aqueous solution into 6 to 10 volumes of glacial acetic acid. The products so obtained were of high purity and very low ash content (64), but were somewhat depolymerized as evidenced by a lowered viscosity and a lower optic rotation.

The hyaluronate preparations sometimes contain nitrogenous impurities, glycogen, and organic and inorganic sulfate. Nitrogenous impurities shown by an increased ratio of total N to hexosamine N can be removed by repeated adsorption. Glycogen, present especially in skin and umbilical cord preparations, is indicated by too low analytical figures (see below) and is removed by incubation for a few hours with either filtered saliva or a commercial amylase. Sulfates are difficult to remove, especially in umbilical cord hyaluronate. The S content may be lowered by dissolving the purified material in 10 per cent calcium acetate or chloride and fractionating with 25 to 33 per cent ethanol. After the mixture has stood at 0° for 24 hours the bulk of the sulfate is found in the precipitate.

Several authors have reported simplified methods for the preparation of hyaluronate. One such method (18) yielded a preparation which after removal of some insoluble material contained only 16 per cent of standard hyaluronate.

The preparations in this laboratory are analyzed for N by the Kjeldahl method, for hexosamine, for acetyl by the Kuhn-Roth method, for uronic acid by the Freudenberg modification of the Tollens-Lefevre method, and for ash.

The rotation $[\alpha]_D^{20}$ of the sodium salt is $-70°$, that of the free acid usually is somewhat lower, depending on the time of contact with acetic acid. The viscosity relative to 0.9 per cent sodium chloride at 37° of a 0.3 per cent solution of the sodium salt usually is approximately 4, and varies with the source. From vitreous humor by fractionation with $(NH_4)_2SO_4$ at a pH of about 9.0, fractions of widely varying viscosities were obtained. This is probably due to enzymatic depolymerization in vivo. The acid is poly-disperse in other sources as well, as pointed out by Blix and Snellman (4), who measured viscosities and streaming double refraction of flow. The latter was positive for both hyaluronate and chondroitin sulfate of cartilage. The particle length of different preparations of hyaluronate varied between 4800A° and 10,000A° with the higher values

in preparations of umbilical cord. The corresponding molecular weights were estimated to be in the order of 200,000 to 500,000 with the assumption of an unbranched chain and a length of 10A° of the basic disaccharide unit. The molecular weights, according to the authors, should be regarded as minimal values.

X-ray measurements of sodium hyaluronate gave only diffuse patterns in powder and fibre diagrams.

The viscosity of isolated hyaluronate is much smaller than that of the fluids from which it is obtained. Whether this is due to oxidative breakdown, as suggested by some authors (83), or to another process of disaggregation or depolymerization cannot be decided with certainty. In the native state hyaluronate forms gels of varying rigidity. For example a cyst formed by an undiagnosed malignant tumor of the bone had the consistency of a 4 per cent agar solution. The cyst content was liquefied rapidly by incubation with testicular hyaluronidase. Samples of human synovial fluid exhibit viscosities ranging from very low values to high viscosities which are outside the range of our viscosimeters. The dilution curve of the tumor fluids showed a sharp drop in viscosity with dilution (55) and the same has been shown by Ragan (72) for normal human joint fluid. Obviously the viscosity in these fluids is of a non-Newtonian type On the other hand, isolated hyaluronate has all the characteristics of a highly asymmetrical molecule. Aqueous solutions have high viscosity, they possess bi-refringence of flow, and can be spun into threads of considerable tensile strength.

The difference between the viscosity of the native fluid and that of isolated hyaluronate may be due to complex formation with accompanying protein, but there is no experimental proof for this. The mobility of isolated hyaluronate in the Tiselius apparatus at pH 7.8 was identical with that of the fast component in tumor and synovial fluids (55, 3). No fraction with a mobility intermediate between protein and polysaccharide was observed. Furthermore, on incubation of the native fluids with hyaluronidase, the viscosities drop rapidly to the values for the proteins present. It is concluded from these experiments that hyaluronate exists in these fluids in varying degree of polymerization and aggregation, the latter being due to weak secondary bonds either between the polysaccharide molecules themselves or between protein and polysaccharide. The weak bonds are broken by the processes involved in the isolation.

The chemical structure of hyaluronic acid is unknown. Equimolar quantities of hexosamine, acetyl and uronic acid have been found by analysis. Glucosamine has been isolated (as the crystalline hydrochloride) in yields up to 95 per cent of the colorimetrically determined amount. The presence of glucuronic acid was established by oxidation of hyaluronic acid by HNO_3 and isolation of saccharic acid as the crystalline acid potassium salt. From an enzymatic digest of synovial hyaluronate with pneumococcus hyaluronidase, glucuronic acid has been isolated as thiosemicarbazone (66). Hyaluronic acid, in contrast to chondroitin sulfate, probably contains little or no branched chains. This is indicated by the almost quantitative hydrolysis by some hyaluronidase preparations. It can be deduced from enzymatic data (see later) that the

basic unit is a disaccharide with a free aldehyde group present in the acetylglucosamine moiety.

The occurrence of ester linkages in the hyaluronate chains has been recently suggested (4), on the basis of the alkali sensitivity of the polysaccharide. In 4–5 days in absence of O_2 at 20° about one milliequivalent of acid was produced for about every 4 disaccharide units. Apparently the reaction came to about the same endpoint with different samples of the polysaccharide and with varying alkali concentrations. About 10 per cent of the weight of the polysaccharide dialyzed through a cellophane membrane and the rest was clearly not of uniform size. The analysis for sodium in the neutral native material did not support the presence of ester linkages. Sensitivity to alkali with production of acidic groups with fission of the chains is common to many other polysaccharides, for example cellulose, though the mechanism is not clearly understood.

Hyaluronate on injection into animals is not antigenic. Attempts to confer antigenicity to hyaluronate have been unsuccessful (29).

II. HYALURONIDASE. A. *Occurrence.* The enzyme hydrolyzing hyaluronic acid, hyaluronidase, was first found in autolysates of a rough type II pneumococcus (58). Its occurrence was shown further in other types of pneumococci both smooth and rough, in strains of hemolytic streptococci of groups A and C, and in some anaerobes (62). Its occurrence in the animal body in extracts of spleen and of ciliary body and iris, has been reported (62).

The very wide distribution of hyaluronidase, however, was found after Chain and Duthie (8) discovered that the hyaluronidase activity of testicular extracts paralleled the activity of the "spreading factor", reported earlier by Duran-Reynals (27) and by McClean (46). Hyaluronidase activity, as measured by the hydrolysis or depolymerization of hyaluronic acid and hyaluronate-containing fluids, was then demonstrated in virulent and avirulent strains of type I pneumococcus (57), in a group C hemolytic streptococcus and in rabbit skin (ibid), in the venoms of many poisonous snakes, like copperhead, Black Tiger (9), rattlesnake (15), in bee venom (9), in Staph. aureus, Cl. Welchii, (9), and others.

B. *Determination of Hyaluronidase.* For the determination of hyaluronidase biological, chemical and physico-chemical methods have been used.

1. *Biological methods.* The spreading reaction has been reviewed by Duran-Reynals (12) and need be mentioned here only briefly. India ink probably is the best indicator, since with it the area of spreading is more circumscribed than with diffusible dyes. The most convenient dye is the blue dye T1824, since it is non-toxic and available in sterilized isotonic solution. An improved quantitative method for the estimation of the spreading reaction has been published by Humphrey (28) who injected the spreading agent intracutaneously into groups of 6 guinea pigs. The animals were killed 20 minutes after injection, the skin was removed immediately and the size of the bleb measured at the inner surface of the skin. The minimal diffusion dose is the least amount of enzyme which will produce a 20 per cent increase of the mean area over the mean area of the controls.

The spreading reaction cannot be considered as an accurate assay of hyaluronidase. Some unspecific agents give spreading reactions. Furthermore, the correlation between spreading reaction and physico-chemical methods of hyaluronidase estimation is poor. (See later.) The influence of pressure on the assay of hyaluronidase by the spreading reaction has been stressed in a recent article (23A). Most investigators however have in the past used a constant and small volume of 0.1 or 0.2 cc. Control injections with identical volumes have probably been used by all investigators. Humphrey has emphasized the importance of the rate of the initial reaction by using a reaction time of only 20 minutes.

The decapsulation of mucoid hemolytic streptococci of groups A and C has been compared with hyaluronidase activity estimated by other methods (47). However, these organisms lose their capsules under various conditions without added enzyme. It is not known whether or not this is due to hyaluronidase produced by the same cultures (see later). The decapsulation of streptococci can hardly be considered as a convenient method for hyaluronidase estimation.

2. *Physico-chemical methods.* Three methods will be discussed: *a*, the mucin clot prevention test (abbreviated M.C.P.); *b*, the viscosity reducing method, and *c*, the turbidimetric method.

a. This method is based on the observation that native hyaluronic acid in acid solution coprecipitates with protein in a typical fibrous "mucin" clot. After incubation with hyaluronidase the quantity of the clot is reduced and the character of the precipitate changes from a fibrous to a flocculent precipitate, until finally the solutions remain clear. The test was apparently first used by Robertson, Ropes and Bauer (76) and has been modified by McClean and his collaborators (52).

The test is simple and suited for serial determinations and is probably the most sensitive test for low concentrations of hyaluronidase. However, it is positive only with crude hyaluronate, the mechanism of its action is not clear and it has a very poor correlation with other tests. The M.C.P. test is carried out in the native fluid or the isolated and redissolved protein salt. Pure hyaluronate precipitates protein in acid solution in flocculent form. (The difference between native and isolated hyaluronate has been discussed on p. 337.)

The error of the M.C.P test has been estimated (48) as ±25 per cent. The author found no correlation between M.C.P. units and viscosity reducing units, nor a constant trend with the potency or the source of the enzyme. Similarly in another paper from the same laboratory (77) the ratios of M.C.P. over viscosity reducing units with different micro-organisms varied between 1:20 to 1:162. The reason for the discrepancy between the two tests cannot be entirely the existence of different hyaluronidases in different sources (see later).

b. The viscosity reducing method. Advantages of this method are its accuracy and its simple kinetics. It has the disadvantage of being cumbersome and not easily carried out in serial experiments. The first systematic use of this method was published by Madinaveitia and Quibell (41) with a crude extract of acetone precipitated vitreous humor as substrate. The time required to reach half

viscosity was found to be inversely proportional to the concentration of enzyme. Details as to substrate and sodium chloride concentrations, pH and temperature vary somewhat in different laboratories. The procedure used in this institute is as follows: 5 cc of a 0.3 per cent solution of pure sodium hyaluronate in 0.1 mol acetate buffer of pH 6.0 containing 0.15 mol sodium chloride are warmed to 37° for about 5 minutes. One cc. of enzyme dissolved in the above buffer mixture is added, and 5 cc. of the mixture are immediately transferred to an Ostwald viscosimeter in a constant temperature bath of 37.0°. The viscosities are determined repeatedly, until less than half viscosity of the mixture is reached. One unit is defined as the amount of enzyme required to reach half viscosity in 30 minutes. The viscosity of the substrate mixture in the control is about 4.0 relative to the solvent. The viscosity of the substrate remains constant for at least 2 weeks at 4° without preservative.

c. The turbidimetric method. This method, described first by Kass and Seastone (34), is based on the observation that pure hyaluronate at pH 4.2 gives a fairly stable colloidal suspension with dilute serum, whereas depolymerized hyaluronate remains clear. The method is reproducible to about 10 per cent, requires little substrate and is readily applied in large series. In a modification developed in this laboratory (21) the time of incubation is kept constant; in the original procedure the time is varied, the enzyme concentration being constant. In this modification the reaction is carried out in the presence of sodium chloride. Dilution of the enzyme in 1 cc. volumes of 0.1 M acetate buffer of pH 6.0 is warmed to 37° for 5 minutes. To each tube 1 cc. of the substrate is added, containing 4 mgm. per cc. of sodium hyaluronate in M/10 acetate of pH 6.0 and 0.3 mol sodium chloride. The mixtures are incubated for 30 minutes at 37°, immersed in a 60° water bath for 10 minutes to inactivate the enzyme, and 1 cc. aliquots are pipetted into test tubes containing 3.0 cc. of 0.5 M acetate of pH 4.2 and 1 cc. of acidified horse serum and mixed. The serum solution is prepared by 10 fold dilution of the sterile normal horse serum (containing no preservative) with 0.5 mol acetate buffer of pH 4.2 and acidifying to pH 3.1 with 4 N hydrochloric acid. The turbidities are read after 30 minutes at room temperature in a Coleman spectrophotometer using a wavelength of 580 mμ. One unit is defined as the amount of enzyme which reduces the turbidity given by 0.2 mgm. of hyaluronate to that given by 0.1 mgm. The values are read from a standard curve. The latter has been checked with 10 different preparations of hyaluronate from different sources. All points fall on the same S shaped curve. In the region between 0.05 and 0.15 mgm. the curve is almost a straight line.

The ratio of the units determined viscosimetrically to those determined turbidimetrically with 4 different hyaluronidase preparations varied only from 1.2 to 1.4. In this comparison the conditions of the two tests were identical with the exception of the substrate concentration.

3. *Chemical methods.* Chemically hyaluronidase activity has been determined by measuring the increase in reducing sugar, or by the increase in liberated acetylglucosamine. In both methods pure hyaluronate of known hexosamine and uronic acid content should be used. The main disadvantage of the reducto-

metric method is its complicated kinetics, obviously the resultant of the action of different enzymes, which simultaneously split glucosidic linkages (see later). In this laboratory the method has been used extensively in the identification of hyaluronate from different sources. Either the Hagedorn-Jensen or the ceric sulfate method in the modification of Miller and Van Slyke was used and the results calculated as per cent of total reducing sugar expressed as equivalents of glucose. The substrate concentration usually was 0.5 per cent

The increase in acetylhexosamine (68) has been used by some investigators (9) using synthetic acetylglucosamine as standard. It was pointed out that this method with the pneumococcal hyaluronidase indicated freeing of acetylhexosamine in excess of 6 times of the total weight of polysaccharide present (57). A similar result was reported by Hahn for testicular and Cl. perfringens hyaluronidase (19). Whether the measured rate of the color reaction of the enzymatic hydrolysis over that of synthetic N-acetylglucosamine is due to a structural difference in the acetylglucosamine or due to the linkage to the uronic acid in the hydrolysis product has not been decided.

In a recent paper Humphrey (31) confirmed the discrepancy between reducing and acetylhexosamine values on hydrolysis of hyaluronate with enzymes of bull testis, Cl. Welchii, and two strains of streptococci. The excess of the acetylhexosamine varied with different enzyme preparations and was not as high as found in this laboratory with pneumococcal enzyme. He further reported that the enzymic hydrolysates gave a direct colour in acid solution with Ehrlich's reagent without preliminary treatment with alkali, which is necessary to form an oxazoline derivative from acetylglucosamine. The author discusses whether oxazoline rings are preformed in the hyaluronate molecule. The alternative that enzymatically formed acetylglucosamine is labile and easily condenses to a heterocyclic compound seems more probable. The hydrolysis of chondroitin sulfate by testicular enzyme showed no excess of acetylhexosamine over the reducing values (32).

The method obviously cannot be used for the determination of hyaluronidase activity.

C. *The Mechanism of Hyaluronate Degradation by Hyaluronidases.* The dual nature of the two glucosidic linkages in hyaluronic acid, one belonging to the N-acetylglucosamine, the other to the glucuronic acid moiety, suggests that the depolymerization and the hydrolysis into monosaccharides require two enzymes. A comparison of hyaluronidases of different origin, measured viscosimetrically and reductometrically, indicated that hyaluronidases were mixtures of at least two enzymes, one attacking the long chain molecules, the other hydrolysing the aldobionic acid units formed (57). This conclusion was based on the observation that pneumococcal hyaluronidase hydrolysed the substrate almost to 100 per cent of the theoretical amount whereas testicular hyaluronidase, which showed a much faster rate in the viscosimetric tests than the bacterial enzyme, hydrolyzed the substrate to only approximately 50 per cent. The testicular enzyme on prolonged hydrolysis exceeded the 50 per cent reduction, but the total reduction was considerably short of 100 per cent. The addition of pneumococcal

hyaluronidase to the non-hydrolysed residue brought about complete hydrolysis, while addition of fresh testicular enzyme had a negligible effect.

A still greater discrepancy between reductometric and viscosimetric activity was found with leech enzyme (60). One preparation as a flavianate reached a maximum hydrolysis, measured reductometrically of 40 per cent of the theory. Yet the same enzyme preparation turbidimetrically assayed 330 units, that is, it was equal to the best testicular preparations we prepared. This preparation showed an interesting difference from other hyaluronidases in the turbidity test, being uninfluenced by the absence of sodium chloride. Hahn (19) working with a purified enzyme prepared from aqueous extracts of the heads of leeches reported a maximal hydrolysis of only 26 per cent of the theory. Other preparations of leech enzyme in this laboratory analysed reductometrically as high as 70 per cent of the theory. It seems quite obvious that in purification one activity is increased at the cost of the other. Another difference between two hyaluronidases, one from bull testes, the other from streptococci, has been reported by Rogers (78). This author investigated the stimulation of hyaluronidase production by streptococci on addition of hyaluronate to a simplified culture medium. If instead of hyaluronate a digest of hyaluronate obtained with streptococcal hyaluronidase, was added to the culture medium, no stimulation of hyaluronidase production took place. If, however, a digest of hyaluronate obtained with testicular hyaluronidase was added, the stimulation of hyaluronidase production was almost as great as with native hyaluronate. Furthermore, the diffusible fraction of the latter hydrolysate was inactive, while the activity remained in the non-diffusible fraction (unfortunately the hyaluronate in these experiments was rather impure). Recently, similar observations were reported on the hyaluronidase production of Cl. Welchii (79).

Other evidence for the existence of several enzymes was brought forward by Hahn (20) who claims to have separated two enzymes from testicular extracts. One enzyme, called mucopolysaccharidase (abbreviated M.P.) hydrolyses to the aldobionic acid stage, the other muco-oligosaccharidase, (abbreviated M.O.) hydrolyses the disaccharide to monosaccharides. Since the M.P. liberates acetylglucosamine during hydrolysis, the free aldehyde group must belong to the hexosamine moiety, therefore the M.O. was called a glucuronidase. The two enzymes were fractionated by $(NH_4)_2SO_4$ and lead acetate, the M.O. activity precipitating in the first fractions. M.P. activity was measured by a viscosimetric method, and M.O. activity by a reductometric method. The substrate of the latter was hyaluronate previously digested for 4 days with M.P. followed by dialysis, the concentrated non-diffusible residue being the substrate. There was no linear relation between the percentage of hydrolysis and the enzyme concentration. The total reducing values must have been quite small, since the substrate concentration was very low, only 0.5 mgm. glucose equivalent per cc., and a turnover of only 10 per cent was measured.

To this reviewer it seems highly desirable to test the presence of the M.O. with a well defined substrate. One would expect the disaccharide formed by the 4 day hydrolysis to be diffusible. There also remains the question of whether the M.O.

can or cannot attack the glucuronidic linkage in the native hyaluronate. The hydrolysis of the linkage in the native chain presumably would also cause a decrease in viscosity, while the end groups with the two enzymes ought to be different.

One distinct enzyme, β-acetylglucosaminidase, present in crude testicular and in crotalus hyaluronidase, as well as in sources showing no hyaluronidase activity, does not seem to be involved in hyaluronidase activity (13). This enzyme was found by Helferich and Iloff (24) in emulsin of sweet almonds. The substrate used was β-phenylacetylglucosamine. The enzyme was separated from testicular hyaluronidase with $(NH_4)_2SO_4$, lead acetate or $CuSO_4$ or adsorption on kaolin. Hahn (20) likewise has reported good evidence for the separation of hyaluronidase from β-glucosaminidase activity.

In summary, hyaluronidases of different origin undoubtedly contain different enzymes, but their number and mode of action are not fully known.

D. *Influence of Environmental Factors on Hyaluronidase Activity.* 1. *Effect of pH.* The effect of pH on the rate of hyaluronidase activity varies with the source of the enzyme, the salt concentration and the test method. In the reductometric procedure testicular hyaluronidase has a double optimum, one at about pH 4.5, the other at about pH 5.7 which has been interpreted as another indication for the presence of two distinct enzymes (57). The pH optimum of pneumococcus and Cl. Welchii hyaluronidase was found in the reductometric procedure at 5.8 (62). By a viscosimetric method, McClean (48) found a marked dependency on the salt concentration of the activity at different pH levels. In M/60 buffer the optimum of testicular hyaluronidase was at pH 6.8, while in M/6 buffer, the optimum shifted to pH 5.0. A similar behaviour was found with Cl. Welchii enzyme. Moreover, at the pH optimum the activity of the enzyme was much greater in the more dilute buffer. The marked difference found may have been partly due to differences in chloride ion concentrations (see later), especially with hyaluronate of umbilical cord. With the turbidimetric method, the pH optimum of testicular hyaluronidase in M/10 acetate in presence of 0.15 M sodium chloride was found to be 6.0 (21).

2. *Effect of salts.* Robertson et al. (76) reported that dialyzed Cl. perfringens hyaluronidase had no effect on a dialyzed "synovial mucin." Addition of phosphate or other salts in increasing concentrations (up to 0.1 M) resulted in increasing activity. Madinaveitia and Quibell (42) reported a marked influence of salts, especially sodium chloride, on the activity of testicular hyaluronidase, as determined viscosimetrically. The optimum was between 0.07 and 0.17 mol sodium chloride.

In this laboratory the influence of salts, especially sodium chloride, was found to be much more marked with the turbidimetric and viscosimetric methods than with the reductometric. The same behaviour had been shown with lysozyme and its substrate (59). The influence of sodium chloride on the rate of depolymerization depended further on the source of the substrate. This is illustrated in the following table in which the same enzyme preparation, highly purified testicular hyaluronidase, has been tested.

It can be seen from this table that the apparent activity of a sample varied from 100 to 167 units per mgm. The wide variation in apparent activity in the absence of NaCl should be noted. By the quantitative turbidimetric test the different hyaluronates precipitated an equal amount of serum protein. Analytically, they were identical within the limits of the errors of the methods. Quantitative determinations in 100 mgm. of the tumor preparations after ashing with HNO_3 showed only a trace of chloride. The chloride effect and the differences between the hyaluronate preparations seem to be due to the presence of competitive inhibitors, as described in the next chapter. It is evident from these experiments that the purity and origin of the substrates greatly influence the titres of enzyme preparations.

3. *Inhibitors of hyaluronidases.* McClean (47) reported the inhibitory action of heparin and chondroitin sulfate on the in vitro decapsulation of streptococci by testicular hyaluronidase. Three-tenths per cent of the pure NH_4 salt of heparin suppressed the decapsulation of an enzyme concentration equivalent to one viscosity reducing unit. From 0.001 to 0.01 per cent of the NH_4 salt further inhibited the action of one unit of hyaluronidase on 0.1 per cent hyaluronate. Hyaluronate partly depolymerized by precipitation with acetic acid likewise

TABLE 1

SOURCE OF HYALURONATE	NO NaCl; UNITS/MGM.	0.15 M NaCl; UNITS/MGM.
Umbilical cord	5	100
Tumor	67	167
Synovial fluid	67	100
Strep. C.	25	125

inhibited the decapsulation, while hyaluronate depolymerized enzymatically was without inhibitory effect. Gastric mucin also inhibited the decapsulation, but Shiga-Kruse polysaccharide and a blood group A hapten had no effect. It should be noted that all inhibitors are or contain acid polysaccharides; the two neutral polysaccharides tested were without action.

In this laboratory the inhibiting activity of heparin on the hydrolysis of hyaluronate was tested and results are shown in table 2. Heparin desulfurated with oxalic acid-barium oxalate had no effect. The inhibition with chondroitin sulfate was considerably less than with heparin, the ratio between the two being approximately 1:100.

The inhibitory action of normal human and rabbit serum on pneumococci, Cl. Welchii and streptococcus hyaluronidase seems to have been reported first by this laboratory (26). The reductometrically determined activity of the first two was inhibited 25 to 50 per cent by normal sera, while with streptococcus hyaluronidase the inhibition was smaller. McClean (47) reported the inhibition of guinea pig, rabbit, sheep, horse, mouse and human serum on hyaluronidase prepared from bull, rabbit and mouse testes, from streptococcus, Cl. Welchii, and from viper and scorpion venom. The inhibitory activity of the sera from

different species against any one enzyme showed considerable variation. The inhibitory action of the sera was associated with the pseudoglobulin fraction. McClean suggested that the inhibition was due to some polysaccharide fraction of the serum.

Meyer et al. (62) found the hydrolysis of free hyaluronate twice as fast as that of an equimolar concentration of a protein salt of the acid, prepared from synovial fluid. In experiments of Kass and Seastone (34) the reduction of the turbidity of a preformed horse serum-polysaccharide precipitate required 10 times the enzyme concentration than that of the polysaccharide alone. More recently Leonard and Kurzrok (36) reported that normal sera inhibit the liberation of tubal ova by testicular hyaluronidase (see later).

The inhibitory action of serum on hyaluronidase of animal and bacterial origin recently was the object of three papers (18). A complete discussion of all the experimental evidence presented and the conclusions drawn from them seems

TABLE 2

	UNITS/MGM.	% INACTIVATION
Hyaluronidase + Hyaluronic acid	83	
Hyaluronidase + Hyaluronic acid + NaCl	200	
Hyaluronidase + Hyaluronic acid + 0.040 mgm. heparin	3	96
Hyaluronidase + Hyaluronic acid + 0.040 mgm. heparin + NaCl	200	0
Hyaluronidase + Hyaluronic acid + 0.004 mgm. heparin	12	85
Hyaluronidase + Hyaluronic acid + 0.0004 mgm. heparin	40	52
Hyaluronidase + Hyaluronic acid + 0.00004 mgm. heparin	83	0

impossible in this review and does not seem warranted. The invasive agent, hyaluronidase, is counteracted by enzymes of the sera of different species, ranging from man to carp, which were said to destroy hyaluronidase. Crude bacterial and some animal hyaluronidases in turn contain other invasive enzymes, which counteract the defensive enzymes of the sera, which in turn are again counteracted by another serum enzyme and so forth.

Experimentally, hyaluronidase was incubated with serum in a salt buffer mixture and the solution mixed with a very impure hyaluronate of umbilical cord in the presence of phosphate. The hyaluronidase activity was measured viscosimetrically. Phosphate was found by Haas to inhibit the action of serum on hyaluronidase, while other salts, especially borate, promote it. The enzymatic nature of the inhibiting agent, in serum was concluded from the temperature sensitivity, the dependence on pH and the inactivation by heating. This evidence appears inconclusive. The activity of sera among comparable data in different tables is very variable. The rate of inactivation was reported to depend

considerably on the quality of the polysaccharide (18, p. 78). (A sample of hyaluronate prepared according to Haas in this laboratory contained 16 per cent of sodium hyaluronate and over 50 per cent inorganic material.) Furthermore, the inhibition by human sera depends not only on concentration but likewise on purity of the hyaluronidase (21). That is, the protein impurities present in hyaluronidase likewise react with the serum. The inactivation by sera is not only prevented by phosphate, but at least in part reversed on further incubation. The inhibitory action of serum furthermore is so dependent on the milieu that it seems very difficult to draw conclusions from experiments in which the complex systems are further complicated by adding crude bacterial filtrates or snake venoms. To this reviewer the data presented do not seem essentially different from the inhibition by heparin and other substances, except that the inhibitors in serum are more thermolabile. The great dependency on milieu conditions is interpreted by this reviewer as being due to competitive reactions between various proteins among themselves and for the acid substrate. That hyaluronidase can be active in the presence of serum proteins is shown not only by its spreading reaction in the skin, but its action in pathological human joint fluid and other fluids and its in vivo activity on sedimentation rate of blood when injected intravenously.

4. *Antisera to hyaluronidases.* The inhibition of the spreading reaction by specific antisera has been discussed by Duran-Reynals (12). Various papers published since then support the previous conclusions that antisera to hyaluronidases suppress the in vivo and in vitro activities of hyaluronidases, but that they are strictly specific for the source of the enzyme. Thus McClean (48) finds no crossing over with antisera against Cl. Welchii to Cl. septicum. Streptococcal enzymes were found to be group, but not type specific. An antiserum against crude bovine testicular hyaluronidase neutralized bovine, but not mouse testis hyaluronidase. Highly purified testicular hyaluronidase, however, has never been obtained in antigenic form (49).

E. *Bacterial Hyaluronidases.* Increase in hyaluronidase production by the addition of hyaluronate to the medium was first reported by McClean and Hale (49) for Cl. Welchii. The stimulating effect of added hyaluronate was studied systematically by Rogers (77, 78). In groups A and C hemolytic streptococci and in two strains of Cl. Welchii enzyme production was proportional to the concentration of hyaluronate added and rose from a titre of 1:50 M.C.P. units to as high as 1:100000 units, when the pH was maintained at neutrality by 8.5 per cent sodium glycerophosphate. One of two strains of staph. aureus produced a titre of 1:3200 units, which was not increased by hyaluronate, the other had a low titre with or without added hyaluronate. The titre was not raised either in a strain of Cl. septicum (77, 78). In staph. aureus the hyaluronidase production was raised by addition of a papain digest of peptone to the medium, while in Cl. septicum it could be increased by some specially prepared peptone (79).

A systematic study of hyaluronidase production in relation to serological groups and types, to capsule formation and to the sources of the strains in hemoly-

tic streptococci was published by Crowley (11). A total of 376 strains were tested, 308 of which belonged to group A. Of the latter, 64 strains, all belonging to types 4 and 22, showed hyaluronidase activity. Seventy-two and 96 per cent respectively of all types 4 or 22 examined were hyaluronidase producers. The titres varied from 50 to 1000 M.C.P. units/cc. Thirty-five out of 55 group C strains and all of 13 examined of group G strains of a variety of types showed hyaluronidase activity. All hyaluronidase positive strains were non-capsulated, but many of non-capsulated strains were hyaluronidase negative. Little correlation was found between hyaluronidase production and virulence. Of streptococcus strains isolated from 127 cases of scarlet fever, 27.6 per cent; of 136 cases of sore throat, 36.8 per cent showed hyaluronidase activity. Of the latter half of the strains belonged to groups C and G. Twenty and four-tenths per cent of hyaluronidase positive strains were isolated from wounds, burns or impetigo contagiosa and 23 per cent from healthy throats. Of 6 cases of puerperal fever, 3 were hyaluronidase producers, among them a group C strain isolated from the blood of a fatal case. However, in another case where the organism was isolated from the blood stream, it was capsulated and negative for hyaluronidase. In some cases highly contagious strains were isolated from contacts in different individuals. The organisms were capsulated and none produced hyaluronidase. The author concludes that there was no evidence for correlation of hyaluronidase production of streptococci with any particular type of infection, nor with virulence of streptococci for man.

Hyaluronidase production of streptococci as reported by Crowley merits some discussion, especially in regard to negative findings in non-encapsulated strains. It was reported from this laboratory (57) that streptococcal hyaluronidase, present in the medium as well as in purified preparations, showed a rapidly diminishing activity as compared to pneumococcal hyaluronidase. Furthermore, the same strains of both group A and C organisms varied greatly in hyaluronidase production from completely negative to high concentrations, although grown under apparently identical conditions. In animals the culture medium and purified samples gave a pronounced spreading reaction regardless of the in vitro titre of hyaluronidase activity. It was suggested that the enzyme may be inactivated, the inactivation being reversible in vivo, while no reversibility could be demonstrated in vitro. Hale (22) confirmed the anomalous behaviour of streptococcal hyaluronidase in the viscosimetric tests but only when carried out at pH 4.6. Furthermore, exposure of the enzymes of groups A and C streptococci to pH 4.6 before mixing with the substrate resulted in complete inactivation of the enzyme. At neutrality the enzyme showed a normal fall in viscosity proportional to the enzyme concentration. In our experiments the substrate was the sterile pleura fluid of a patient with mesenthelioma. The reaction was carried out at a pH of about 7.0. The reductometric tests were carried out at pH 5.8 with pure hyaluronic acid as the substrate. Inactivation of the enzymes was apparent with both methods. Whether or not the inactivation is enzymatic has not been determined.

In view of these experiences with streptococci it may be questioned whether

failure to demonstrate hyaluronidase is synonymous with absence of the enzyme. A renewed study of hyaluronidase in streptococci seems indicated, not only in regard to hyaluronidase production, but also in regard to spreading activity, which in our experience is not confined to types 4 and 22 of group A organisms.

The correlation between capsule formation and hyaluronic acid production has been studied by Seastone (82) in a group C streptococcus. The evidence there given for the rôle of the hyaluronate capsule in the virulence of the infecting organism has been corroborated by Hirst (25), who protected mice by the intraperitoneal injection of leech extracts against fatal infection with a strain of group C streptococcus, while against a group A strain only feeble protection resulted. Blundell (5) obtained a greater mean survival time in mice infected with a group A organism, when treated with testicular hyaluronidase than without treatment. McClean (47) reported failure to protect either against groups A or C organisms with testicular hyaluronidase. Kass and Seastone (34) protected mice against 10 to 100 M.L.D of groups A and C organisms by repeated injections of testicular hyaluronidase. No protection resulted from injection of inactivated hyaluronidase. The specificity for streptococcus was demonstrated by the failure to protect against the fatal infection with a type I pneumococcus. The failure of other authors to obtain this effect was explained by the time lag between injections of the enzyme, which allowed the organisms to regenerate their capsules. Kass and Seastone further studied the in vitro effect of hyaluronidase on streptococci. Phagocytosis of group A organisms by human leucocytes and the killing power of whole blood was greatly increased by incubation of the organisms with hyaluronidase. In controls with pneumococci no change resulted from incubation with hyaluronidase.

In *pneumococci* no correlation between the amount of hyaluronidase produced and the clinical virulence, or between enzyme production and the type of organism, was found (30). Type I organisms rarely produced the enzyme.

In *staphylococcus aureus* no correlation between enzyme production and virulence was reported by Boe (6). However, Schwabacher et al. (81) found among 654 coagulase positive strains of staphylococci 86.7 per cent producing both α-lysin and hyaluronidase, 6.9 per cent hyaluronidase positive and α-lysin negative, 4.4 per cent hyaluronidase negative and α-lysin positive and 2 per cent negative for both. None of the 160 coagulase negative strains produced either α-hemolysin or hyaluronidase. The author concludes that clinical virulence is associated fairly clearly with α-hemolysin production and to some extent with hyaluronidase production. What part hyaluronidase plays in determining the virulence of a strain of Staphylococcus was not clear.

In Cl. Welchii, McClean, Rogers and Williams (52) found 12 out of 32 strains produced hyaluronidase. Of these 12 strains 11 produced toxin and 10 were derived from clinical cases of gas gangrene. In Cl. septicum 20 out of 20 strains in vitro produced hyaluronidase. Of these 4 were derived from clinical cases of gas gangrene, 9 were stock laboratory strains and 7 were contaminants of "healthy" wounds. Seven out of 15 strains of Cl. oedematiens produced a low

titre of hyaluronidase. Six of these were stock laboratory strains and one was derived from a toxemia in sheep. McClean et al. (52) proposed the determination of hyaluronidase and lecithinase for early diagnosis of gas gangrene in wound infections. Guinea pigs were infected intramuscularly with organisms of the gas gangrene group and the presence of the enzymes was determined in the muscle extracts and edema fluid. With Cl. Welchii as infective agent the muscle contained 2 M.C.P. units 2 hours after infection. With Cl. septicum the infection developed more slowly than with Cl. Welchii. Edema fluid was not obtained before 12 hours after infection, when it as well as the muscle did contain hyaluronidase in low concentrations. With Cl. oedematiens the infection proceeded still more slowly than with Cl. septicum. In some instances 12 hours after infection hyaluronidase was detected in the edema fluid, but since many strains did not produce hyaluronidase in vitro, this organism was the least satisfactory from the point of view of detection of enzymes in the body fluids. The authors further proposed to utilize specific antisera for the differential diagnosis of the infecting agents in combination with the enzyme determinations, as in positive tests only the specific antisera will inhibit the enzyme of the infecting agent (50).

Clinically the proposal of McClean et al. was tested by MacLennan (39) in 39 cases, among them 12 cases of clinical gas gangrene, the others from various infections or from non-infected wounds. Only in one case, infected with a streptococcus pyogenes, was a positive M.C.P. test obtained.

McClean and Rogers (51) ascribed the failure of MacLennan at least in part to the treatment of the clinical cases with antisera. In guinea pigs the administration of antisera inhibited and sometimes suppressed the appearance of hyaluronidase and lecithinase. They further tested whether the size of the infecting dose explained the difference between the animal experiments and clinical cases. A minimal infecting dose of organisms did not diminish the titre in the infected tissues of the guinea pig. They further studied a combination of infecting agents. Cl. histolyticum, which produced no hyaluronidase, caused death in combination with a streptococcus earlier than Cl. histolyticum alone. The strain of streptococcus used did not produce illness in the guinea pig either in presence or in absence of added hyaluronidase. A combination of Cl. histolyticum with the streptococcus and added hyaluronidase caused death even earlier than in the previous experiments.

MacLennan and associates (40) maintained in a later article that the determination of hyaluronidase and lecithinase was unsatisfactory as a diagnostic procedure.

F. *Substrates of Hyaluronidase.* The difference in the apparent activity of testicular hyaluronidase on hyaluronate of different sources has been discussed above. Highly purified testicular hyaluronidase has been stated to hydrolyze beside hyaluronate only two other mucopolysaccharides, one a monosulfuric acid ester of hyaluronic acid obtained from cornea (54), the other the chondroitin sulfate of hyaline cartilage (57) (43). The hydrolysis of chondroitin sulfate is

of considerable interest, since compounds of the general composition of this acid occur beside hyaluronate, in mesodermal tissue, like umbilical cord (64) and skin (56) in concentrations about equal to that of hyaluronate. The chondroitin sulfate of skin differed in rotation from that of cartilage (56). The two acids were found, in collaboration with Dr. Z. Dische, to differ further in their resistence to alkali and considerably in the time-intensity curve of their reaction with carbazole. Moreover, chondroitin sulfate of skin, and to a lesser degree of umbilical cord, in contrast to that of cartilage were resistant to hydrolysis with testicular hyaluronidase.

Is chondroitin sulfate of cartilage a substrate of hyaluronidase, or is it hydrolysed by another enzyme copresent in the testicular preparations? The evidence points to the latter explanation. Chondroitin sulfate is not attacked by pneumococcal or streptococcal or leech hyaluronidases. While it might be argued that the bacterial enzymes are to a lesser degree true mucopolysaccharidases, this cannot be claimed for the leech enzyme. Furthermore, purified testicular hyaluronidase after precipitation by acetone lost most of its chondroitin sulfate hydrolysing power, while the hydrolysis of hyaluronate was hardly impaired. However, it seems remarkable that hyaluronidase activity towards hyaluronate runs practically parallel with the activity towards chondroitin sulfate, whether the fractionation is carried out with metal salts or with adsorption. The enzyme hydrolyses chondroitin sulfate with the production of a disaccharide without hydrolysing the sulfate linkage. Humphrey in a recent paper (32) found no hydrolysis of chondroitin sulfate with streptococcal enzyme and a negligible effect with an enzyme prepared from Cl. Welchii. Crude and a highly purified testicular enzyme however at pH 4.8 and 6.0 liberated reducing sugars with chondroitin sulfate at a rate comparable to hyaluronate. Some bacterial extracts have been reported to hydrolyse chondroitin sulfate by splitting off inorganic sulfate (69).

III. Hyaluronic acid and hyaluronidase in animal physiology and pathology. A. *Ocular Fluids.* Hyaluronic acid has been isolated from vitreous humor of cattle, swine, and sheep and from the aqueous humor of cattle. Its concentration in the normal fluids has been determined by analysis of the hexosamine content. This procedure involves only a small error, since the protein, being very low in these fluids, contributes very little to the total hexosamine. The concentration of hyaluronic acid found by isolation was in fair agreement with that calculated from the hexosamine values. The polysaccharide content of vitreous humor varies considerably with the species, from a low of about 9 mgm. per cent in the cat to a high of about 48 mgm. per cent in cattle eyes, while the concentrations in aqueous humor showed little variation (67).

The origin of the polysaccharide in the ocular fluids is not known with certainty. While the permanent structures of the vitreous body seem to be of retinal origin (44), the fluid part probably originates mainly in the ciliary epithelium which (in this country at least) is considered a secretory epithelium. The presence of hyaluronate in the fluids and its absence from serum have been cited as evidence for its origin from a secretion rather than a dialysate (65). Hyaluronate on injection into the anterior chamber of rabbits disappears rather

rapidly. After paracentesis of the anterior chamber the acid reappears (as judged by analysis of the hexosamine) after the return of the protein content to normal. It was concluded that the acid was constantly produced and removed, presumably after enzymatic hydrolysis. Hyaluronidase has been reported in extracts of the ciliary body and iris (67) though Chain and Duthie (9) denied its occurrence in these tissues and suggested bacterial contamination as a source of the enzyme. The presence of hyaluronidase has been definitely established in pooled aqueous humor obtained by aspiration of the eyes of freshly killed cattle. The aqueous humor was lyophilised and the residue dissolved in 1/10 of the original volume. Four units (turbidimetric) of hyaluronidase were found per cc. of the concentrate, corresponding to 0.36 unit per cc. of the native fluid (21). It is of considerable interest that the hyaluronate present in the same sample was found to be almost completely depolymerized. While the vitreous humor of the same eyes had a hyaluronate concentration (turbidimetric) of 0.75 mgm. per cc., the lyophilized aqueous humor showed only 0.00015 mgm. per cc. By hexosamine estimation the hyaluronate concentration of the vitreous humor ranged from 0.32 to 0.52 mgm. per gram (the lower figures are probably explained by the difference in specific gravity and evaporation of water). In aqueous humor the hyaluronate concentration was about 0.03 mgm. per gram, a value in fair agreement with the concentration found by isolation of an impure fraction (0.046 mgm. per gram). Since the concentration by the turbidimetric method amounted only to 0.0015 mgm. per cc., we may conclude that the hyaluronate in aqueous humor is depolymerized to about 95 per cent of the total. As this depolymerization presumably occurred in vivo, we assume that hyaluronate is after enzymatic depolymerization constantly removed through the normal exit channels of the eye.

What part hyaluronate plays in the physiology and pathology of vitreous humor is unknown. It probably takes part in the maintenance of the turgor of the vitreous body. The effect of intravitreal injections of purified testicular hyaluronidase in rabbits is being studied by Dr. L. von Sallmann. Injections of about 100 turbidimetric units appear to liquefy the vitreous humor in vivo. The cause of simple glaucoma may well be explained by inhibition of hyaluronidase in the eye.

B. *Synovial Fluid.* Bauer et al. (1) have reviewed the physiology of synovial fluid. The view expressed in their paper that synovial fluid contains a glycoprotein instead of a dissociated mucopolysaccharide seems no longer tenable (see p. 337). Furthermore, the concept of the origin of synovial fluid as a dialysate to which the "mucin" is added by the passage of the fluid through the connective tissue does not seem probable. If such a mechanism existed then pleura and peritoneal fluid and lymph should likewise contain the "mucin" which is not demonstrably the case. From the appearance of a viscous fluid in tissue culture of synovial tissue (86) it may be concluded that hyaluronate is a secretory product of some cells of the synovial lining. This seems to be borne out by the isolation of hyaluronic acid from a synovioma, not only at the original site of the tumor, but in metastases in the liver (60).

The concentration of hyaluronic acid in normal human synovial fluid has not

been determined by isolation (see below). In cattle synovial fluid obtained by aspiration of astragalo-tibial joints the concentration by isolation was found between 20 to 25 mgm. per cent (66). From knees of patients with rheumatoid arthritis the acid was isolated in 3 different samples in concentrations of 60, 132 and 206 mgm. per cent. However, it may be questioned whether cattle synovial fluid as collected, is normal and comparable to the fluid in human joints. Its viscosity is low compared to that of normal human fluid taken at autopsy or from living patients with no pathology of the joints (72). By extrapolation of the dilution curve of human synovial fluid the relative viscosity of the latter was estimated as greater than 180, whereas that of the cattle fluid was less than 10. The volume of the fluid of cattle taken immediately after slaughter is quite large, sometimes over 50 cc. It seems possible that due to long standing and trauma of the animals before killing the cattle fluid is actually a mixture of synovial and edema fluid.

The hyaluronic acid and protein contents and the viscosities of about 30 pathological fluids have been determined (74). The viscosity was variable and never higher than that of normal knee joints. The protein content was over 3 per cent, that of normal human and cattle fluid below 3 per cent. The viscosity practically was all due to hyaluronate, since on incubation with enzyme the relative viscosity fell to slightly above 1.0, corresponding to the viscosity of the remaining protein solution (66).

By the turbidimetric method the hyaluronic acid content varied from 80 to 270 mgm. per cent. No direct proportionality between viscosity and hyaluronate concentration was found. In normal human and in cattle synovial fluid the hyaluronic acid concentration could not be determined turbidimetrically by the above method. In contrast to the pathological fluids the normal fluids in the turbidimetric procedure are precipitated as a coherent fibrous clot while the pathological fluids with few exceptions precipitated in the form of a stable turbidity, as do vitreous humor and the tumor fluids. On incubation of the normal fluids with 0.01 unit of enzyme, far too little to decrease the apparent hyaluronate concentration, the clot formation of the normal fluids is prevented and hyaluronic acid can be determined turbidimetrically. The amount found varied between 80 and 150 mgm. per cent as compared to 80 to 270 mgm. per cent in pathological fluids. That is, both the concentration and especially the total amount of the fluid in pathological joints is larger than that found in normal joints. Hyaluronidase in contrast to normal ocular fluid could not be demonstrated in synovial fluid, although the absence of fibre formation may suggest its presence in low concentrations in pathological joint fluids.

C. *Hyaluronic Acid in Tissues.* The presence of hyaluronate in tissues when isolation is not possible, can be made probable enzymatically by the use of a variety of bacterial and animal enzymes. Judging from the failure of isolation experiments and the absence of any effect on capillary permeability by purified hyaluronidase, hyaluronic acid is not present in capillary cement. The presence of hyaluronic acid in all connective tissues has not been established. Positive spreading reactions in stomach and intestinal walls, in the uterus, in striated

muscle, fasciae and tendons with crude testicular extracts have been reported (12, 14). It remains to be determined whether the spreading is due to depolymerization of hyaluronate or chondroitin sulfate. In calves' tendon the polysaccharide is a sulfate ester and is hydrolyzed by testicular hyaluronidase at a fast rate (60). Only in skin and in umbilical cord do larger quantities of hyaluronate occur. It seems remarkable that in both tissues chondroitin sulfate is present in a concentration roughly equal to that of hyaluronate. Both substances have to be regarded as constituting essential components of the interfibrillar ground or cement substances in these tissues. It seems of great importance to gain more information about the nature of such cement substances especially for the understanding of the mechanism of diseases of the mesenchymal tissues as well as of wound healing and the problem of ageing.

Beside hyaluronate and chondroitin sulfate the mucopolysaccharide isolated from amyloid tissue may likewise be such a cement substance. This substance is a monosulfuric acid ester, containing acetylhexosamine and a uronic acid in equimolar proportion. The hexosamine has been isolated and characterized as d-glucosamine (60). The behaviour of the mucopolysaccharide towards hyaluronidases shows that it is not derived from hyaluronic acid. From its enzymatic behaviour and its reactions with alkali and with carbazole it appears to be closely related to heparin. It may well be possible that this substance accumulates in excessive amounts in certain pathological conditions, while in small concentrations it may be a normal component of some mesenchymal tissues.

Histologically the presence of hyaluronic acid has not been demonstrated, since no staining method was known. Hale (23) recently proposed a histochemical test for hyaluronic acid. After a specific fixation of the tissue, slices are treated with iron hydroxide. The iron combines with the acid polysaccharides. The bound iron is stained as Prussian blue. For the differentiation between hyaluronate and other polysaccharides incubation of the tissue slices with streptococcal hyaluronidase is recommended, since this enzyme does not attack chondroitin sulfate. No experimental results with this method have been reported.

Highly polymerized hyaluronate in a concentration of about 1 per cent shows in smears typical metachromatic staining with toluidine blue, while lower concentrations failed to show the effect. The metachromasia is prevented on incubation with hyaluronidase (unpublished experiments). The concentrations present in tissues are probably never high enough to contribute to the metachromatic staining. Therefore, the metachromasia of tissues, done under appropriate conditions, probably does indicate sulfate containing compounds (37) in connective tissue probably largely chondroitin sulfate. The viscous material demonstrated by Bensley (2) in regenerating and young tissue and by Maximow (45) in tissue culture, possibly indicates hyaluronate. The rôle of calcium and ascorbic acid in the cement substances of plants and animals has been discussed by Reid (74A).

Robb Smith (75) has studied histologically in vitro the effect of testicular and streptococcal extracts on the muscle of guinea pigs. Reticulin or collagen

were reported not to be affected, but the reticulin membranes became separated from the muscle fibres. The author suggests that there may be a layer of mucopolysaccharide between the reticulin membrane and the sarcolemma. The changes were absent in heated extracts and were prevented by streptococcus antisera when streptococcal hyaluronidase was employed.

A tentative picture of the development of cement substances in mesodermal tissue combining the admittedly scanty data may be presented as follows: The young fibroblast secretes into the surrounding tissue spaces hyaluronic acid, a precursor of collagen, and a chondroitin sulfate. By local acidification in the immediate neighborhood of the cells the first fibres are produced by the polysaccharides from the native soluble collagen, which denature into the insoluble fibre, on the surface of which lies a sheet of the polysaccharides. With ageing of the fibres, the polysaccharide layer becomes thinner and the hyaluronate is replaced more and more by chondroitin sulfate. Only in metabolically very active connective tissue like that of skin, hyaluronic acid production continues in appreciable quantities. The rôle of ascorbic acid (87) in the process of fibre formation may be that it actually is a component of chondroitin sulfate, replacing in the chain some of the glucuronic acid molecules.

D. *Hyaluronic Acid and Erythrocyte Sedimentation Rate.* Hyaluronic acid in common with other asymmetrical molecules increases the erythocyte sedimentation rate in vitro or after intravenous injections (61). It was further observed in this laboratory that purified testicular hyaluronidase decreased in vitro the sedimentation rate of blood of patients, especially in rheumatic fever. This action was first thought to be due to the hyaluronidase contained in the enzyme preparations. However, it became apparent that hyaluronidase activity as determined by chemical or physico-chemical methods was not proportional to the action on sedimentation rate. This became still more obvious when it was found that enzyme preparations fractionated by lead acetate no longer acted on sedimentation rate, when lead was removed by Na_2S instead of dialysis, while the hyaluronidase activity was not altered. In further experiments by Dr. C. Ragan (unpublished) the mechanism of the action on erythrocytes was found to be due to spherocyte formation, caused by an enzymatic action of unknown nature on the erythrocyte membrane.

The effect of hyaluronate on erythrocytes seems comparable to the precipitation of large particles, such as hemocyanine, liver particles and a number of viruses by some colloids of high molecular weight (10). The precipitation occurs at neutrality and apparently is not due to polar forces. Depolymerized hyaluronate had no effect. It seems remarkable that hyaluronate of high polymerization had a much larger effect quantitatively than either chondroitin sulfate or heparin (see fig. 1 of Cohen and personal communication). The mechanism of this action does not seem clear, but may be connected with complex coacervation (7) and obviously deserves further study.

E. *Hyaluronidase in the Animal Body.* The rôle of testicular hyaluronidase in testis probably is confined to the dispersion of the cumulus cells in the process

of fertilization (53) (80) (15A). Apparently only mature spermatozoa contain a high concentration of the enzyme. Testes of birds, amphibia and reptiles contain little (12). The mucus of the cervical plug is not acted upon by hyaluronidase (53) (80). The hyaluronidase of human semen has been investigated by several workers (32A) (35A) (86A). Kurzrok et al. by a modified M.C.P. test found a critical value for hyaluronidase of 50 million sperm per cc. Hyaluronidase activity did not run parallel to sperm population. Some apparently normal semen was found to be deficient in hyaluronidase, concomitant with inability to cause fertilization. Mixing of semen and testicular hyaluronidase or application of hyaluronidase in powder form to the cervix was said to have increased considerably the number of successful fertilizations.

No conclusive data are available for the occurrence of hyaluronidase in blood. Intravenously injected enzyme is rapidly eliminated from the circulation.

The demonstration of hyaluronidase in aqueous humor has already been mentioned.

A relatively large concentration of hyaluronidase occurs in skin (57). In press juice of ground rabbit skin very little activity can be demonstrated, but when the material is autolysed in presence of toluol in phosphate buffer of pH 5.0, a relatively large concentration is found in the supernatant solution. Chain and Duthie (9) were unable to find any hyaluronidase in skin. However, the authors likewise failed to obtain a spreading reaction with skin extracts, in contrast to Duran-Reynals' (12) and our experiments. The higher concentration in autolysed material may be due either to the liberation of enzyme bound to cell structures or conversion of inactive into active material. The latter hypothesis seems more likely. Of interest in this connection is the observation of Duran-Reynals that washed glandular tissues of poisonous snakes are almost free of hyaluronidase, whereas it appears with the glandular secretion.

The concentration of potential hyaluronidase in skin may be still higher than that found, since during autolysis considerable destruction of added enzyme occurs.

The occurrence of hyaluronic acid and hyaluronidase in skin may point to a rather rapid turnover of the former. This turnover may be regulated by certain hormones. Luria and Zappasodi (38) found a significant increase in the spreading effect in rabbits treated with luteinizing hormone. Other workers (84) have found a decreased reaction under the influence of follicular hormone. The edema fluid of the sex skin of monkeys in the estrus phase has been reported to be similar to synovial fluid (9). The turgescence of the sex skin of the baboon during the menstrual cycles was studied by Clarke (9A). A considerable part of the water uptake of the skin was apparently bound water since no free edema fluid could be demonstrated. The deturgescence at the end of the estrus cycle supplied the animal with fluid for 11 days. We may assume that this fluid was bound by hyaluronic acid.

Localized hyaluronidase action has been suggested as a possible explanation

for the bullae formation in pemphigus (38A). However, after intradermal injection of hyaluronidase into normal people and into patients suffering from a variety of rheumatic diseases, bullae were never observed.

F. *Clinical Application.* The use of hyaluronidase in artificial insemination has been mentioned above.

On injection of purified testicular hyaluronidase into the knee joints of patients suffering from rheumatoid arthritis the viscosity of the fluid was remarkably reduced. No change was observed in the underlying disease. One week after the injection, the viscosity of the fluid had returned to the original level (73).

In one patient having a mesenthelioma of the pleura and peritoneum purified testicular hyaluronidase of bull or ram was injected intraperitoneally to facilitate removal of a fluid of honey-like consistency (85). Without hyaluronidase injection, evacuation of the fluid was difficult and incomplete. After injection of 16,000 to 80,000 units of hyaluronidase, fluid of low viscosity could be completely removed in a short time. No immediate harmful effects of the injections were apparent. Injections and paracentesis were repeated at various intervals. The tumor finally did not seem to produce as much fluid as originally. On autopsy large tumor masses were found in peritoneum and pleura with little fluid mostly encapsulated.

A connection between hyaluronic acid and hyaluronidase and rheumatic diseases, especially rheumatic fever, has probably been suspected by many workers. In a number of articles Guerra (16) recently reported the inhibiting effect of sodium salicylate on the spreading of India ink, in rabbits injected with crude testicular hyaluronidase. The area of spreading without hyaluronidase was reduced by 20 per cent with 0.07 and by 31 per cent with 0.10 gram per kgm. of sodium salicylate. With hyaluronidase injection the area of spreading was reduced 57 per cent with 0.07 gram per kgm. and 66 per cent with 0.1 gram per kgm. of sodium salicylate. Sulfadiazine did not decrease the spreading effect of hyaluronidase.

In human subjects intradermal injection of the enzyme with the dye T1824 was said (17) to cause, in individuals with active or inactive rheumatic fever, unique reactions with enormous diffusion of the dye and local edema that sometimes involved the entire arm. Salicylate inhibited the spreading reaction in those cases. The type of reaction was also observed in one male with exanthematic typhus.

In this laboratory salicylate in vitro in equivalent or higher concentration has been found to be without effect on the depolymerization or hydrolysis of hyaluronate. However the marked depressing action on skin diffusion in rabbits has been confirmed (J. A. Coss, personal communication). The explanation of this effect as well as the beneficial action of salicylate on some rheumatic manifestations may be found in an inhibition of hyaluronate production of mesenchymal cells.

CONCLUSIONS

It is obvious from this review that the relationship between bacterial infection and the hyaluronidase system has continued to receive more of the attention

of investigators than has the metabolism of hyaluronic acid and its rôle in animal physiology and pathology. From the scanty data available, it seems obvious that the functions of the skin, of the ocular fluids, of synovial fluid and of the connective tissues in general must depend in part on the quantity and degree of aggregation of hyaluronic acid. Gels formed by the acid serve partly as the cement which holds cells together. In other structures as in the joint they protect internal surfaces, or they are part of the viscous barriers as in some connective tissues which regulate the exchange of metabolites and water. Thus the physiological aspects of hyaluronic acid as well as of other mesodermal cement substances seem to be of even greater importance than their rôle in infection.

REFERENCES

(1) Bauer, W., M. W. Ropes and H. Waine. Physiol. Rev. **20**: 272, 1940.
(2) Bensley, S. H. Anat. Record **60**: 93, 1934.
(3) Blix, G. Acta Physiol. Scand. **1**: 29, 1940.
(4) Blix, G. and O. Snellman. Ark. f. Kemi., Mineral. o. Geol. **19A**: no. 32, 1945.
(5) Blundell, G. P. Yale J. Biol. **14**: 373, 1942.
(6) Bøe, J. Acta Path. Microbiol. Scand. **21**: 587, 1944.
(7) Bungenberg de Jong, H. G. and W. A. L. Dekker. Kolloid Beihefte **43**: 213, 1936.
(8) Chain, E. and E. S. Duthie. Nature **144**: 977, 1939.
(9) Chain, E. and E. S. Duthie. Brit. J. Exper. Path. **21**: 324, 1940.
(9A) Clarke, R. W. Am. J. Physiol. **131**: 325, 1941.
(10) Cohen, S. J. Biol. Chem. **144**: 353, 1942.
(11) Crowley, N. J. Path. and Bact. **56**: 27, 1944.
(12) Duran-Reynals, F. Bact. Rev. **6**: 197, 1942.
(13) East, M. E., J. Madinaveitia and A. R. Todd. Biochem. J. **35**: 872, 1941.
(14) Favilli, G. Sperimentale Arch. biol. **89**: 724, 1935.
(15) Favilli, G. Nature **145**: 866, 1940.
(15A) Fekete, E. and F. Duran-Reynals. Proc. Soc. Exper. Biol. and Med. **52**: 119, 1943.
(15B) Glick, D. Ann. Rev. Biochem. **11**: 51, 1942.
(16) Guerra, F. J. Pharmacol. and Exper. Therap. **87**: 193, 1946. Arch. del Inst. de Cardiol. de Mex. **16**: 3, 1946.
(17) Guerra, F. Science **103**: 686, 1946.
(18) Haas, E. J. Biol. Chem. **163**: 63, 89, 101, 1946.
(19) Hahn, L. Ark. f. Kemi., Mineral. o. Geol. **19A**: no. 33, 1945.
(20) Hahn, L. Ark. f. Kemi., Mineral. o. Geol. **21A**: no. 1, 1945.
(21) Hahnel, E. and K. Meyer. Unpublished experiments.
(22) Hale, C. W. Biochem. J. **38**: 368, 1944.
(23) Hale, C. W. Nature **157**: 802, 1946.
(23A) Hechter, O. Science **104**: 409, 1946.
(24) Helferich, B. and A. Iloff. Ztschr. Physiol. Chem. **221**: 252, 1933.
(25) Hirst, G. K. J. Exper. Med. **73**: 493, 1941.
(26) Hobby, G. L., M. H. Dawson, K. Meyer and E. Chaffee. J. Exper. Med. **73**: 109, 1941.
(27) Hoffmann, D. C. and F. Duran-Reynals. J. Exper. Med. **53**: 387, 1931.
(28) Humphrey, J. H. Biochem. J. **37**: 177, 1943.
(29) Humphrey, J. H. Biochem. J. **37**: 460, 1943.
(30) Humphrey, J. H. J. Path. and Bact. **56**: 273, 1944.
(31) Humphrey, J. H. Biochem. J. **40**: 435, 1946.
(32) Humphrey, J. H. Biochem. J. **40**: 442, 1946.
(32A) Joel, C. A. and E. Eichenberger. Schweiz. Med. Wchnschr. **75**: 601, 1945.

(33) Kabat, E. A. J. Biol. Chem. **130**: 143, 1939.
(34) Kass, E. H. and C. V. Seastone. J. Exper. Med. **79**: 319, 1944.
(35) Kendall, F. E., M. Heidelberger and M. H. Dawson. J. Biol. Chem. **118**: 61, 1937.
(35A) Kurzrok, R., S. L. Leonard and H. Conrad. Am. J. Med. **1**: 491, 1946.
(36) Leonard, S. L. and R. Kurzrok. Endocrinology **37**: 171, 1945.
(37) Lison, L. Histochimie animale, Méthodes et problèmes. Paris, 1936.
(38) Lurie, M. B. and P. Zappasodi. Arch. Path. **34**: 151, 1942.
(38A) MacCardle, R. C., J. P. Baumberger and W. C. Herold. Arch. Derm. and Syph. **47**: 517, 1943.
(39) MacLennan, J. D. Lancet p. 433, 1944.
(40) MacLennan, J. D. and R. G. MacFarlane. Lancet **2**: 301, 328, 1945.
(41) Madinaveitia, J. and T. H. H. Quibell. Biochem. J. **34**: 625, 1940.
(42) Madinaveitia, J. and T. H. H. Quibell. Biochem. J. **35**: 453, 1941.
(43) Madinaveitia, J. and M. Stacey. Biochem. J. **38**: 413, 1944.
(44) Mann, I. C. The development of the human eye. London, 1928.
(44A) Mann, T. and C. Lutwak-Mann. Ann. Rev. Biochem. **13**: 25, 1944.
(45) Maximow, A. Proc. Soc. Exper. Biol. and Med. **25**: 439, 1927.
(46) McClean, D. J. Path. and Bact. **34**: 459, 1931.
(47) McClean, D. J. Path. and Bact. **54**: 284, 1942.
(48) McClean, D. Biochem. J. **37**: 169, 1942.
(49) McClean, D. and C. W. Hale. Biochem. J. **35**: 159, 1941.
(50) McClean, D. and H. J. Rogers. Lancet **1**: 707, 1943.
(51) McClean, D. and H. J. Rogers. Lancet **2**: 435, 1944.
(52) McClean, D., H. J. Rogers and B. W. Williams. Lancet **1**: 355. 1943.
(53) McClean, D. and I. W. Rowlands. Nature **150**: 627, 1942.
(54) Meyer, K. and E. Chaffee. Am. J. Ophth. **23**: 1320, 1940.
(55) Meyer, K. and E. Chaffee. J. Biol. Chem. **133**: 83, 1940.
(56) Meyer, K. and E. Chaffee. J. Biol. Chem. **138**: 491, 1941.
(57) Meyer, K., E. Chaffee, G. L. Hobby and M. H. Dawson. J. Exper. Med. **73**: 309, 1941.
(58) Meyer, K., R. Dubos and E. M. Smyth. Proc. Soc. Exper. Biol. and Med. **34**: 816 1936. J. Biol. Chem. **118**: 71, 1937
(59) Meyer, K. and E. Hahnel. J. Biol. Chem. **163**: 723, 1946.
(60) Meyer, K. and E. Hahnel. Unpublished experiments.
(61) Meyer, K., E. Hahnel and R. R. Feiner. Proc. Soc. Exper. Biol. and Med. **58**: 36, 1945.
(62) Meyer, K., G. L. Hobby, E. Chaffee and M. H. Dawson. J. Exper. Med. **71**: 137, 1940.
(63) Meyer, K. and J. W. Palmer. J. Biol. Chem. **107**: 629, 1934.
(64) Meyer, K. and J. W. Palmer. J. Biol. Chem. **114**: 689, 1936.
(65) Meyer, K. and J. W. Palmer. Am. J. Ophth. **19**: 859, 1936.
(66) Meyer, K., E. M. Smyth and M. H. Dawson. J. Biol. Chem. **128**: 319, 1939.
(67) Meyer, K., E. M. Smyth and E. Gallardo. Am. J. Ophth. **21**: 1083, 1938.
(68) Morgan, W. T. J. and L. A. Elson. Biochem. J. **28**: 988, 1934.
(69) Neuberg, C. and W. M. Cahill. Enzymologia **1**: 22, 1936.
(70) Palmer, J. W., E. M. Smyth and K. Meyer. J. Biol. Chem. **119**: 491, 1937.
(71) Pirie, A. Brit. J. Exper. Path. **23**: 277, 1942.
(72) Ragan, C. Proc. Soc. Exper. Biol. and Med., in press.
(73) Ragan, C. and A. De La Mater. Proc. Soc. Exper. Biol. and Med. **50**: 349, 1942.
(74) Ragan, C., E. Hahnel and K. Meyer. Unpublished experiments.
(74A) Reid, M. E. Physiol. Rev. **23**: 76, 1943.
(75) Robb-Smith, A. H. T. Lancet **2**: 362, 1945.
(76) Robertson, W. van B., M. W. Ropes and W. Bauer. J. Biol. Chem. **133**: 261, 1940.

(77) ROGERS, H. J. J. Path. and Bact. 56: 284, 1944.
(78) ROGERS, H. J. Biochem. J. 39: 435, 1945.
(79) ROGERS, H. J. Biochem. J. 40: 583, 1946.
(80) ROWLANDS, I. W. Nature 154: 332, 1944.
(81) SCHWABACHER, H., A. C. CUNLIFFE, R. E. O. WILLIAMS AND G. J. HARPER. Brit. J. Exper. Path. 26: 124, 1945.
(82) SEASTONE, C. V. J. Exper. Med. 70: 347, 1939.
(83) SKANSE, B. AND L. SUNDBLAD. Acta. Physiol Scand. 6: 37, 1943.
(84) SPRUNT, D. H. AND S. McDEARMAN. Endocrinology 27: 893, 1940.
(85) STEWART, W. A. AND K. MEYER. Unpublished experiments.
(86) VAUBEL, E. J. Exper. Med. 58: 63, 85, 1933.
(86A) WERTHESSEN, N. T., S. BERMAN, B. E. GREENBERG AND S. L. GARGILL. Am. J. Urol. 54: 565, 1945.
(87) WOLBACH, S. B. Am. J. Path. 9: 689, 1933.

PATHOGENETIC FACTORS AND PATHOLOGICAL CONSEQUENCES OF DECOMPRESSION SICKNESS

H. R. CATCHPOLE[1] AND ISIDORE GERSH[1]

Naval Medical Research Institute, Bethesda, Md., and Department of Anatomy, The Johns Hopkins University, Baltimore, Md.

For almost fifty years no comprehensive review of the pathological consequences of decompression sickness has appeared. The reviews of Paul Bert (1878) and of Heller, Mager and von Schrötter (1900) defined the etiology of the disease and formulated the main aspects of its pathological results. The views of these authors have remained essentially unchallenged. But since that time a great clarification of the physical factors involved in aeroembolism has come about and a need has arisen to integrate more recent work in pathology with the older literature, with the objective of reinterpreting both in terms of these physical factors.

In dealing with this literature, key references have been preferred to exhaustive quotation. Deliberate limitations in subject matter have caused the exclusion from discussion of changes referable to the air-containing cavities, where air expansion *per se* gives rise to the pathological effect (pain, rupture, hemorrhage); thus removed from consideration are: aero-otitis media, sinus pain, pain and other effects arising from trapped gases in the intestinal tract, and, as possibly connected with the same phenomenon, pain associated with the teeth (86, 114, 129).

Some of these factors as pertaining to aviation have been dealt with by Armstrong (1). Also excluded are considerations of pathological physiological changes, e.g., in renal, cardiac and metabolic functions, in the gastro-intestinal tract, in the composition of the blood, in blood vessels and in the general phenomenon of shock. Effects due to anesthesia have been omitted, but a section on drugs and exercise is included. Effects directly referable to acute and chronic anoxia have also been omitted. In so doing, an effort has been made to separate the results of low oxygen on the body as a whole from those traceable to aeroembolism.

This review falls into two sections. In the first, physical and mathematical considerations governing the uptake and elimination of gases by the body and by individual tissues are discussed, and the conditions are defined for the relative susceptibility of a given site to bubble formation. In the second section, the gross and microscopical pathology of organs and tissues following decompression are related to these factors. While the effects of pressure and those of altitude are usually considered separately, the three categories of caisson disease, decompression sickness of divers, and aeroembolism of aviators, are regarded as basically similar entities, in which pathological differences are, or will prove to be, adequately explained by reference to the physical factors involved.

[1] Present address: Department of Pathology, University of Illinois, College of Medicine, Chicago, Ill.

PATHOGENETIC FACTORS IN DECOMPRESSION SICKNESS. *Physical Considerations. Mathematical principles.* The phenomena accompanying gas bubble evolution in the bodies of animals subjected to decompression have led recently to a revival of interest in the purely physical mechanisms involved in the separation of a gaseous phase from a saturated or supersaturated solution of a gas. This topic had already been explored in the eighteen seventies (for early literature, see 38, 72). The conditions for stability of a bubble of gas immersed in a liquid saturated with gas at 1 atmosphere have been defined (38, 72, 102) by the relation:

$$P = H + 2\sigma/r$$

where P = pressure in the bubble in excess of atmospheric, H = the hydrostatic pressure at the bubble level, σ = surface tension of the liquid, r = radius of the bubble.

As the radius r of the bubble diminishes, the second term of the above expression gets larger, and when r approaches the dimensions of a water molecule, the excess pressure due to the surface tension of the liquid is counted in thousands of atmospheres (38). There exists, in fact, a lower critical value for r (that is, a minimum bubble size) below which this excess pressure will literally squeeze the gas back into solution (72, 102, 128). Above the minimum size, however, the bubble will grow by diffusion as long as the tension of dissolved gas in its vicinity is greater than the gas tension in the bubble. For a free bubble to arise *de novo* it is necessary that this minimal size be achieved by sufficient molecules simultaneously attaining enough energy to overcome the forces of attraction between them. This was considered to be within the realms of statistical probability by Piccard (128). However, other theoretical treatments and experimental work reported by Dean (38) and Harvey (72) lead to the conclusion that bubbles do not tend to form in liquids spontaneously unless high negative pressures or considerable degrees of superheat are applied.

Gas masses existing in cracks or attached to irregular surfaces, on the other hand, have very different conditions for stability depending on the geometry of the surface, the shape of the gas-liquid-solid junction, contact angles, surface tension and ΔP (see below) which have been worked out for certain situations (72). Such gas masses may be stable at or below a critical size, but they grow indefinitely by diffusion above this critical size; an important property is that their gas content may be increased gradually by successive boostings. Gas nuclei are defined (72) as small invisible masses of gas usually, but not always, attached to a surface, which grow by inward diffusion of gas from the surrounding liquids. These enlarged masses may eventually become detached as free bubbles, or bubbles may become detached leaving behind a nucleus for the growth of other bubbles. The origin of these gas collections is obscure. The condition that they normally be attached to a surface, or contained in a crack, arises from the instability of small free bubbles noted above. Gas monolayers are believed not to promote bubble formation, but the possibility that multilayers of gas may act as nuclei has been suggested (38). Hydrophobic surfaces hold gases very tenaciously (72) but there is no evidence that tissue surfaces possess such properties.

Gas nuclei may be produced in surface cracks or acute angled cavities by statistical fluctuations of gas molecules (72, cf. also 128).

Harvey (71, 72) has introduced a useful expression to define the tendency, ΔP, of a gas to leave a liquid phase:

$$\Delta P = t - P$$

where t = gas tension in the liquid; P = hydrostatic pressure (this may be positive or negative in sign; the latter condition obtains when a pull is exerted on a liquid). Bubbles may form *de novo* when a very high negative pressure is produced in a liquid (when P is negative, ΔP increases). Acting in the vicinity of a gas nucleus, local negative pressures will favor the diffusion of gases into the nucleus. Such "cavitation" of a liquid may arise through pressure pulses, sound waves, Bernoulli effects (motion of liquids through constricted tubes), turbulent motion and through stretching. Increase of t will produce the same effect by providing a richer population of gas molecules for diffusion. Dean (38) attributed most cases of bubble formation in liquids in motion to vortices produced by turbulent flow. However, the streamlining of the vascular flow militates against this view as applied to bubble formation in animals (72).

Models of systems containing gas nuclei were studied by Pease et al. (126). Capric acid (M. P. 31°C) cooled to the point of crystallization provided foci for cavitation in aqueous solutions, and the process was reversed on warming; i.e., gas nuclei were apparently created and destroyed by these procedures. The same authors found that stearate monolayers on glass promoted cavitation; alcohols, amino acids and proteins abolished this effect. Removal of gas nuclei from liquids was early described by Tomlinson (cited by Dean, 38), and since that time bubble nuclei have been variously removed by boiling or partial evacuation of water, followed by standing (38); by centrifugalizing, filtering, boiling or subjecting to 1000 atmospheres pressure (72); and by pressure, evacuation or chemical agents (126). The remarkable properties of liquids freed from gas nuclei in resisting superheat and negative pressures have been frequently described (72, 97, 98). Harvey et al. (73) distinguished between macronuclei, removable from water by centrifugalization, and micronuclei, removable only by high pressures (1090 atmospheres). Freshly drawn blood is completely free of macro- and micronuclei; such nuclei are neither present in the blood plasma, nor attached to formed elements or to any other constituent of the blood. It must therefore be assumed that they are attached to the linings of blood vessels, and that bubbles formed at the sites of the nuclei are released into the circulation (72).

Factors in the growth of gas nuclei in the body. If the concept of $\Delta P = t - P$ be adopted as a measure of the bubble forming tendency, and the presence of gas nuclei in the body be accepted, a number of conditions favoring, or tending to prevent, bubble formation become intelligible. Local production of CO_2 increases t and therefore increases ΔP, and so conduces to bubble appearance (69, 110); muscular activity, besides increasing the CO_2 tension locally gives rise to mechanical tensions and to consequent decrease in P; for both reasons it favors

bubble formation (28, 68, 110, 177). Tissue manipulation of almost any kind (stretching, cutting, crushing) leads to the same result (73). Bone fractures are a potent site of bubble formation (17, 70), since they lead to sharp momentary falls in the value of P. Bubbles form more readily in veins (low P) than in arteries (high P) (72, 177). Among factors delaying or preventing bubble formation in cats stimulated to muscular activity at altitude were: pre-stimulation (109, 177), anoxia (109, 69), traumatization of legs by skinning (109); these treatments promote local hyperemia or hyperventilation, and result generally in an increased value for P combined with faster elimination of nitrogen. Not susceptible of such simple analysis are the effects of vasoconstriction, which are equivocal, and of vasodilatation (aminophyllin), which are negative (137). It is evident that for the body as a whole, ΔP represents an aggregate of pressure differences of constantly varying magnitude in different sites. The situation is summarized by Harvey (72): "at ground level ΔP is nearly zero in tissues to negative in arteries. On rapid ascent to high altitude it is at first positive everywhere, but quickly becomes negative in arteries (due to blood pressure and rapid equilibration with alveolar air), remaining locally high in small vessels (due to CO_2 and N_2 of tissues) for a time, finally becoming zero to negative except for regions of fat deposits or very poor circulation."

Establishment of critical pressure differences. Decompression establishes ΔP values necessary for bubbles to grow. Boycott, Damant and Haldane (21) found experimentally in goats and man that decompression from 2.3 atmospheres to one atmosphere never produced symptoms. On the basis of equivalent gas volumes they argued that a drop of approximately one-half of the original gas pressure would be safe whatever its value (e.g., from 4 to 2 atmospheres, from 6 to 3 atmospheres, etc.). But as will be seen in a succeeding section, this assumption is untenable, and Haldane himself (65) has stated that above six atmospheres it is no longer quite safe to halve the initial pressure. Behnke (5) proceeded on the basis that the difference between tissue gas pressure and the external pressure should at no time exceed 1.3 atmospheres (2.3 minus 1.0 atmosphere). He then calculated decompression rates that would hold the external pressure at not more than 15 lb/sq. in. below the greatest tissue pressure. The views of both Haldane and Behnke have been critically examined by de Burgh Daly and his associates (37, 43). Smaller animals appear to be able to tolerate greater pressure differentials. Guinea pigs survive decompression from 60 lb/sq. in. (gauge) to atmospheric pressure (15 lb/sq.in.), and bubbles are produced with difficulty in rats and mice.

For the occurrence of aeroembolism in man there appears to be a critical altitude of 20,000 to 25,000 feet (0.45 to 0.37 atmosphere) as given by the British authorities (133), or perhaps somewhat higher. Under extreme conditions of stimulation, bubbles appear in cats decompressed to 35,000 feet (0.23 atmosphere). In rabbits, bubbles can not usually be produced at 40,000 feet (0.18 atmosphere); at 45,000 feet they are formed under specified conditions (27). Bubbles scarcely appear in quiet frogs below 60,000 feet, although violent mus-

cular activity reduces the ceiling considerably (177). The smaller rodents again are relatively refractory to bubble formation and rats at altitude, in a state of normal activity, do not yield bubbles at 50,000 feet (177).

Effect of rate of decompression. Since the time of Paul Bert, it has been repeatedly demonstrated that the severity of decompression sickness is directly related to the rate of decompression. Altitude studies by Griffin et al. (62) compared the effect of different rates of ascent (1000 and 5000 ft. per minute) and showed a markedly greater susceptibility to bends after the faster ascents.

Systematic studies on rabbits decompressed to altitude showed a relation of rate of decompression to the incidence and severity of bubble formation (table 1) (27). When decompressed to 45,000 feet in eight minutes or less, mortality was high, and survival time became progressively shorter as the time to altitude decreased. At decompression times exceeding ten minutes to reach 45,000 feet, mortality was low and bubble incidence minimal or zero.

TABLE 1

Effect of rate of decompression on the 50 per cent survival time, the percentage of deaths, and the symptoms of rabbits decompressed to 45,000 ft.

RATE OF DECOMPRESSION	50 PER CENT SURVIVAL TIME	PERCENTAGE OF DEATHS	AVERAGE BUBBLE INCIDENCE SCALE: 0 (NO BUBBLES) TO 5 (MOST NUMEROUS) (27)
	min.		
3–10 sec.	12	81	5 to 2
30 sec. to 1 min.	7	92	3 to 1
2 min. to 5 min.	13	66	2 to 0
6 min. to 8 min.	17	80	1 to 0
10 min.	30	15	0

Similar considerations appeared to control the death of animals receiving injections of air at varying rates into peripheral vessels (140).

Decompression from high pressures and to altitude. The principle that decompression from five atmospheres to one is far more hazardous than from one atmosphere to one-fifth of an atmosphere was demonstrated *in vitro* by Piccard (128). Water saturated with air gave brisk effervescence in the former case, and slow, delayed evolution in the latter. The total *volume* of gas available for release was the same in both instances. The explanation (68, 128) lies in the greater *weight* of gas dissolved at the high pressure, and consequently the greater number of gas molecules available for diffusion into gas nuclei (or for the formation of aggregates of gas by collision). This determines the rapidity of bubble formation, when the pressure is suddenly released. An *in vivo* counterpart of this demonstration is shown by a comparison of bubble frequency and distribution in guinea pigs decompressed from 105 lb/sq. in. gauge pressure to atmospheric pressure, and in rabbits decompressed from ground level to 45,000 feet (51). Bubble formation was far more severe in the former case (table 2) although the relative pressure reduction was seven to one in both. These results are entirely

similar to those reported by de Burgh Daly et al. (37, 43) who compared the results of decompressing rabbits, guinea pigs and rats from 6.3 to 1.0 atmosphere, and from 1.0 to 0.16 atmosphere. It is clear that the reduction in pressure by one half employed by Boycott, Damant and Haldane (21) would actually become more hazardous as the initial pressure increased.

Composition of gas bubbles. It was conclusively shown by Paul Bert (18) that nitrogen formed the major constituent of gas bubbles recovered from animals after decompression from compressed air atmospheres. The rôle played by CO_2 in bubble formation has given rise to some speculation as to the composition of gas bubbles at the site of formation. Gas diffusion constants for the common respiratory gases are approximately equal, and the composition of a bubble will

TABLE 2

Comparison of the distribution of gas bubbles in rabbits decompressed to a simulated altitude of 45,000 feet with guinea pigs decompressed from compressed air at 105 lb. per sq. in. (gauge)

	HIGH PRESSURE	LOW PRESSURE
Extravascular gas bubbles		
Adrenal	Numerous	None
Nerves	Numerous	None
Fat (intracellular)	Numerous	None
Blood vessels		
Arteries and veins	Present	Present
Capillaries	Present	Fat only
Spleen		
Sinusoids	Present	Present
Arteries and veins	Present	Present
Branches of pulmonary arteries	Present in many	Present in few
Liver		
Sinusoids	None	None
Central vein	Few	Few
Liver cells	Watery vacuoles present	No watery vacuoles present
Intestine	Numerous	Few
Muscle	Numerous	Few

therefore be largely controlled by the amount of gases close to it (38). Carbon dioxide is some 50 times more soluble in water than nitrogen. In CO_2 rich regions, such as a contracting muscle, this gas may condition the formation of the primary bubble and represent its principal constituent, at least for a while. Recent experiments (69, 110) support the rôle of CO_2 as a facilitator in bubble formation. When the bubble rich in CO_2 is moved to a body region rich in nitrogen and poor in CO_2, the latter gas will diffuse out and nitrogen in, to give bubbles essentially composed of nitrogen.

Gas Uptake and Elimination. Mathematical principles. Since the first demonstration of the importance of nitrogen in events leading to decompression sickness, attempts have been made to analyse experimentally and describe theoretically the course of gas uptake or elimination when the ambient pressure is raised or

lowered. That gas uptake would present the characteristics of a logarithmic relationship was recognized by Zuntz (185) and by Heller, Mager and von Schrötter (80). Employing the findings of Vernon (168) on the high solubility of nitrogen in fat, Boycott et al. (21) also presented a logarithmic relationship, but indicated a slower rate of saturation per round of circulation than the above authors. They stated further that their computations gave a rough approximation only to the actual rate of saturation for the body as a whole, due to variations in blood flow to the several tissues and to differences in body composition. Experimentally they found in goats a 94 per cent saturation in 3 hours. From this they deduced a similar degree of saturation for man in 5 hours. In computing decompression tables for divers, they arbitrarily assumed the existence of tissues having half saturation times of 5, 10, 20, 40 and 75 minutes, while admitting the possibility of tissues saturating at yet slower rates. Campbell and Hill (23) showed that about one third of the gaseous nitrogen of the human body was removed in the first few minutes while breathing 100 per cent oxygen, but some tissues were found to remain unsaturated after several hours' excess pressure (24). Hawkins, Shilling and Hansen (76) from a study of a large series of experimental dives modified the British decompression tables. They also adhered to the assumption of tissues having half saturation times of 5 to 75 minutes. Decompression times were radically reduced, especially for dives of short duration, in which the "slow" tissues would have become only partially saturated.

That saturation and desaturation curves should be reciprocal was indicated by Boycott et al. (21) and shown experimentally by Shaw et al. (146). The latter further showed that nitrogen absorption obeys Henry's law. A somewhat different concept was introduced by Behnke and his co-workers (5, 6, 7, 8, 12). Total nitrogen of the body was considered to be partitioned between aqueous and fatty phases, and the curve of nitrogen elimination was represented as the sum of two exponential expressions governing respectively the "water" nitrogen and the "fat" nitrogen. There was satisfactory agreement between nitrogen elimination found for man, and that calculated from the equation:

$$Y = \underset{\text{(water)}}{382\,(1 - e^{-0.0085\,t})} + \underset{\text{(fat)}}{458\,(1 - e^{-0.098\,t})}$$

where Y = total elimination of N_2, t = time in minutes, e = base of natural logarithms (12).

Some question was expressed subsequently of the adequacy of the k values cited above. Underwood and Diaz (164) injected radon into the saphenous vein of the dog and studied its elimination through the lung at one minute intervals for four minutes. Part of the gas in the body was shunted to regions of poor circulation and was not measured in the short time interval allowed. They found k values of 0.66 as compared with the much smaller values obtained by Behnke (12), and believed that their figure applied to gas leaving the blood whereas the Behnke values were for gas entering the blood from tissues. This finding was held to resolve the difficulty encountered by Shaw et al. (146) which led to their postulation of a peculiar state of nitrogen supersaturation *in vivo*.

Smith and Morales (150, 151), Morales and Smith (117, 118, 119), Jones et al. (90), and Ferris et al. (48) have contributed further to the analysis of blood-tissue gas exchanges and it becomes necessary to establish the general trend of such attempts to date. It appears to be implicit in the work of Boycott et al. (21) that the total body nitrogen reservoir may be represented as the sum of 5 (or more) arbitrary "tissues" with half saturation times of 5 to 75 minutes. Shaw et al. (146) and Behnke et al. (12) also considered that nitrogen elimination could be represented empirically by one or more exponential equation of the form

$$Y = A\ (1 - e^{-kt})$$

where Y = amount of N_2 eliminated in time t (minutes), A = amount of N_2 originally present, k = constant of elimination, e = base of natural logarithms.

In actual practice two such functions were used and referred to water and fat phases respectively. Underwood and Diaz (164) generalized from the work of these latter authors to the form

$$Q = \Sigma_i Q_0\ (1 - e^{-k_i t})$$

where Q = total amount of gas eliminated, t = time in minutes, $_iQ_0$ = initial amount of gas present in the i state (i.e., in water, fat, etc.), k_i = elimination constant for the i state, e = base of natural logarithms.

Smith and Morales (150) in developing an equation for the uptake of inert gas by tissue regions, e.g., the tissues of a limb, considered the following physical and physiological factors to be operative in gas exchange: blood volume of the region, delivery rate of blood flow, delivery concentration of gas in the blood, tissue volume, gas solubility in each tissue, area of the capillary bed, and tissue permeability to the gas. They derived an expression for tissue regions identical in form with that of Underwood and Diaz (164) but in which the various Q and k values were invested with physiological meaning. These quantities are not to be considered as characteristics of a specific tissue component (water or fat, etc.), but are functions of all tissues of the region, and of the circulation. They concluded (117) that the early, more rapid absorption stage is governed by the blood and by aqueous tissues in close relation to it, and the slower, later stages predominantly by fatty tissues; that these processes are nevertheless simultaneous and conditioned by physical and physiological factors whose importance varies from tissue to tissue and from region to region. These factors may be varied, e.g., by the substitution of one inert gas by another, by increasing blood flow to a region and by altering the degree of fatness. By the method of deriving the general equation employed by these authors, the effects of altering these variables are claimed to become predictable. Further, these quantities may be measured independently of gas uptake, and substitution into the equations developed gave agreements held to be close enough to justify the method of approach (117). Jones et al. (90, cited 117) found curves for nitrogen desaturation of the whole body to conform to the general equation. They emphasize in particular the rôle of the circulation, and believe that all their results are explainable on the basis of gas solubility and blood-tissue perfusion rates.

Gas solubilities. The greater susceptibility of fat individuals to decompression sickness was noted by Boycott and Damant (20) who deplored the fact that increase in experience and technical skill should so often be associated with the increasing waist that accompanies the onset of middle life. This susceptibility was attributed to the greater solubility of nitrogen in fat which becomes a gas reservoir maintaining nitrogen pressure long enough for bubbles to form during the pressure drop of decompression. From this time, considerations of gas solubility in body constituents have occupied an important rôle in the theory and practice of decompression sickness. It is generally held that tissues, with the exception of fat, take up gases in proportion to their water content (23). Data on the solubility of gases in plasma and body fat are incomplete, but their solubility in water and certain oils is better known. Since the effects of salts and proteins (74, 75, 144, 167) and the differences in the nature of the lipids (168)

TABLE 3

Solubilities, relative solubilities and ratio of solubility in oil to solubility in water of respiratory and inert gases presented in order of ascending molecular weight and density. Cited in part from ref. 161

GAS	MOL. WT.	DENSITY AIR = 1.0	SOLUBILITY IN WATER AT 37°C	RELATIVE SOLUBILITY IN WATER He = 1.0	SOLUBILITY IN OIL AT 37°C.	RELATIVE SOLUBILITY IN OIL He = 1.0	RATIO OF SOLUBILITY IN OIL TO SOLUBILITY IN WATER
			cc/100 cc		*cc/100 cc*		
H_2	2.0	0.0695	1.6	1.9	4.5	3.0	2.8
He	4.0	0.138	0.85	1.0	1.5	1.0	1.7
N_2	28.0	0.963	1.3	1.5	6.7	4.5	5.2
O_2	32.0	1.105	2.4	2.8	12	8.0	5.0
A	39.9	1.38	2.6	3.1	14	9.5	5.4
CO_2	44.0	1.529	56.0	66.0	87.6	58.0	1.6
Kr	83.7	2.868	4.5	5.3	43	29	9.0
Xe	131.3	4.525	8.5	10.0	170	110	20.0
Rn	222	7.526	15	17.5	1900	1300	126.0

on gas solubility are relatively small, the values for water and oil may be used as first approximations for plasma and fat (12, 23, 168). Oil and water solubilities of H_2, O_2, CO_2, N_2 and of the inert gases He, Ne, A, Kr, Xe, Rn have been recently critically examined and tabulated (161). Values for some members of the series were newly derived using a radioactive isotope technique, while other values were derived from the older literature. Some of these results are included in table 3.

Nitrogen. If the fat content of the body is 15 to 20 per cent, the fivefold greater solubility of nitrogen in fat compared with water implies that one-half of the total nitrogen of the body is dissolved in fatty tissue (5, 12, 21, 23). During short exposures to excess pressure, fat may act as a reservoir to protect the body against sudden flooding with bubbles (5). Times for essentially complete saturation of the body by nitrogen are given as: man, 5–6 hours (20, 23); dog, 3–4 hours (146); goat, ·3 hours (21). Small animals saturate and desaturate much

more rapidly (21) and require higher pressure differentials to elicit bubble formation.

Oxygen. In atmospheres of compressed air, the amount of dissolved oxygen in tissues represents an inconsiderable part of the dissolved gases. However, the oil-water ratio for oxygen approximates that for nitrogen, and exposure to 100 per cent oxygen under pressure followed by decompression leads to the appearance of gas bubbles (53, 81, 83). For reasons believed to involve tissue damage by oxygen under pressure, bubble formation was of less intensity on decompression from oxygen atmospheres (53). In decompressions to altitude from oxygen at atmospheric pressure the formation of bubbles is greatly reduced, and pretreatment with oxygen is commonly employed for protection under these circumstances (28, 49, 177, 179).

Helium. A low oil-water ratio and the low solubility in both plasma and fat has rendered this gas relatively ideal as a diluent for oxygen in diving (6, 13, 15). The total body content of He at saturation is only 40 per cent that of N_2 and relatively less is present in the "slow" tissues. Time for helium elimination is approximately 50 per cent that required by nitrogen. There is a cutaneous diffusion of helium of 50 cc. per hour which increases rapidly above 28°C with the concurrent sharp rise in peripheral blood flow (14).

Hydrogen. The solubility characteristics of hydrogen are less favorable than those of helium and no special advantage is gained by its use. It has been employed in experimental diving work (183).

Argon, Xenon, Krypton, Radion. The oil/water solubilities of these gases increase roughly in proportion to their atomic weights (91). The last three have found application chiefly as radioactive tracers. Use of argon as a diluent for oxygen in compression studies has served to confirm the view that the production of a highly saturated reservoir of inert gas favors bubble formation. Symptoms of increasing severity were produced by gas mixtures in the following order: oxygen, helium-oxygen, air, argon-oxygen (53), conforming to the oil/water solubilities of helium, nitrogen and argon respectively.

Chemical composition of the body. Arbitrary resolution of the body into lean body mass and fat in studying abnormal pressure effects was suggested by Behnke (6, 9) and a theoretical formulation was developed by Morales et al. (115). The lean body mass, consisting of bone and other tissues, together with essential lipids, maintains in the adult a relatively constant composition. Body density was used as an index of obesity by Behnke et al. (10). By comparing body specific gravity with direct estimations of total fat in the guinea pig, Rathbun and Pace (134) formulated a quantitative relationship between them. A similar relation between specific gravity and fat content was suggested by these authors for man. Computations of body composition from inert gas uptake have been made (6) and methods involving the differential uptake of two inert gases have been suggested as a means of measuring watery and fatty tissues (91, 116, 123) but apparently have not been elaborated experimentally.

Analysis of the effects of decompression on individual tissues requires a breakdown of their composition in terms of fatty and non-fatty components. Values have been derived from the literature for the composition of some tissues of

particular interest in decompression (table 4). The fat content of bone marrow is high (56, 85, 173). Since bone marrow may represent from 2.2 to 4.4 per cent of the body weight, bone marrow fat may comprise up to one-fifth of the total body fat. The adrenal gland is relatively rich in fat (7.5–14 per cent) and in decompression from high pressures a fall in specific gravity of this tissue has been correlated with the presence of extravascular bubbles, paralleling the situation existing in fat tissue proper (52). Brain (156), nerve, liver and muscle are relatively poor in fat; the skin, as a rule, is somewhat richer than these tissues (66).

Tissue blood flow. Inert gas uptake depends among other things on the rate of blood flow, and heavily perfused organs saturate and desaturate rapidly (150). It may be assumed that urine originating in Bowman's capsule would reflect the gas tension of the glomerular blood. Hill and Greenwood (81) early showed experimentally in man that the pressure of nitrogen in the urine became equal to

TABLE 4

Fat content of body tissues and organs

TISSUE OR ORGAN	ANIMAL	FAT CONTENT		REFERENCE
		Per cent wet weight	Per cent dry weight	
Brain	Man	5–8		156
Yellow bone marrow	Pig	87–90		24
	Ox, sheep, horse	90–96		24
	Goat	90–95		24
Red bone marrow	Rabbit	70–90		173
Adrenal gland	Guinea pig	7.5–14 (calculated)	30–54	52
Nerve-sciatic	Guinea pig	5.1 (calc.)	10–24	52
Liver	Guinea pig	1.4 (calc.)	4.9 (range: 0.5–11.8)	52
Muscle	Guinea pig	1.7 (calc.)	6.8 (range: 1.0–16.5)	52
Skin	Rat	6.6–12		66

that in the alveoli within 10 minutes after exposure to increased pressure and that the excess nitrogen rapidly disappeared when the pressure was lowered. Behnke and Yarbrough (15) found that the saturation time for urine in men breathing a He-O_2 mixture was between 30 and 60 minutes, and desaturation time occupied the same interval in the absence of bubble formation. In the latter case, desaturation was delayed. It would be expected that organs with perfusion rates of a comparable order, such as the thyroid and the liver would saturate and desaturate with equal rapidity.

Values for blood flow through tissues and organs have been derived from the literature (table 5). Classical methods of measuring blood flow were used in most instances. However, Jones et al. (90) have calculated blood-thyroid and blood-liver perfusion rates from the exchange of radio-iodine and radio-phosphorus, respectively. Tissues with poor blood supply include fat, bone marrow and resting muscle; active skeletal muscle, heart muscle and brain have a moderate

blood supply and the kidney and thyroid have the richest blood flow. Lying outside this general classification are organs like the spleen with an intermittent circulation and the liver with its dual circulation.

Local peculiarities of the circulation may modify figures based on the total perfusion rate. Gersh (56) contrasted the intermittency and probably reduced rate of flow in splenic sinusoids with the relatively continuous flow in those of the liver. Hepatic blood flow has been estimated to be from one-third to one-seventh of the portal blood flow of the liver (63, 152). In decompression, the slower and presumably more saturated portal blood becomes diluted with less saturated blood from the hepatic artery, thereby lowering the tension of gas available for bubble formation. Organs having blood trapped in sinusoids may on the other hand be exposed to sharp pressure differences between the entering arterial blood

TABLE 5

Blood flow in tissues and organs

TISSUE OR ORGAN	ANIMAL	BLOOD FLOW	REFERENCE
		ml/1000g/min	
Thyroid	Man	4000	90
	Mouse	3240	90
Kidney	Man	3000	90
	Man, dog, rabbit	3400, 3300, 3200	148
Brain	Monkey	850 (range: 600–1000)	40
	Man	650 (range: 460–1040)	142
Heart	Dog	500–750	47
Muscle (dilated)	Dog	57–570	124
(dilated)	Dog	110–580	178
Muscle	Man	24 (range 5-rest to 50-exercised)	90
	Dog	69 (range 30–130)	61
Bone marrow	Goat	26	24
Fat	Man	12–15	90

and the trapped sinusoidal blood, sufficient to produce extravascular or intravascular bubbles. The rapid saturation of all structures of the kidney related to the glomerular apparatus may not extend to portions of the renal tissue having a different vascular pattern.

Capillary density and surface. While descriptive accounts of the capillary distribution in tissues abound, quantitative studies relating capillary surface (S) to unit tissue volume (V) are largely lacking. Values for total and open capillaries of fat tissue were derived (54) from histological studies on injected and frozen-dried preparations respectively. The ratio S/V (cm^{-1}) was computed for fat-rich and fat-poor adipose tissue and compared with estimates for muscular tissue based on data provided by Krogh (101) (table 6). Open capillaries represented from one-half (fat-rich) to one-quarter (fat-poor) of the total capillary area of fat; a similar general relationship exists in muscle (147). For other tissues,

estimates of S/V by indirect means are probably subject to considerable error. Values have been given (147) for the number of red blood corpuscles per cubic millimeter of tissue, which may be regarded as a rough measure of the number of open capillaries. Accepting the fairly reasonable value of 1000 cm^{-1} as a measure of S/V of the capillaries of mouse heart muscle, values for some other tissues have been computed (table 7). Fat-rich adipose tissues have available for gaseous exchange one quarter or less surface than the most poorly supplied muscle, and fat-poor tissues barely equal the poorest muscle. The relative in-

TABLE 6

Capillary surface of tissues in relation to total tissue volume (S/V)

TISSUE	ANIMAL	S/V (cm^{-1})		REFERENCE
		Total (open plus closed capillaries)	Open caps.	
Fat Fat-rich	Rat	51.9	23.5	54
Fat-poor	Rat	222.2	64.1	54
Muscle (a)	Rabbit	190–513		42
(b)	Rabbit	150–295		153
(c)	Dog	494		101
(d)	Guinea pig		186–507	174
(e)	Mouse		486–923	174
Cardiac muscle	Man	1184		175

TABLE 7

Relative values of S/V calculated from the number of red blood corpuscles per c.mm. of tissue as estimated by Sjöstrand (147)

TISSUE	ANIMAL	NO. CORPS./C. MM.	S/V (cm^{-1})
Heart muscle	Mouse	350	1000
Brain-cortex	Mouse	85	240
Cerebellum	Mouse	115	330
Liver	Mouse	280	800
Kidney-cortex	Mouse	300	850
-medulla	Mouse	500	1400
Duodenum	Mouse	160	460
Pancreas	Guinea pig	125	360

adequacy of the capillary bed of fatty tissue is therefore demonstrated quantitatively (54).

Cerebral cortex and cerebellum compare unfavorably with muscle in respect of the capillary bed available for gas transfer. Cobb and Talbot (34) and Cobb (33) concluded that even the most vascular parts of the brain were only one-fifth as vascular as skeletal muscle. Within the central nervous system itself Craigie (35), from measurements of the total length of capillaries in sections of the central nervous system of the cat found the poorest part of the grey matter to be one and one-half times as richly vascularized as the richest part of the white matter. The same was true in the monkey. In the cat Wolff (181) found the

following relative vascularities: parietal cortex, 100; lateral geniculate ganglion, 91; globus pallidus, 58; nerve, 47; and white matter, 43. Thus the white matter is something under one-half as richly vascularized as the grey. In addition, capillaries in the white matter are said to be larger than those of the grey (33) which would make S/V values for the white matter even more unfavorable.

Occupying a somewhat intermediate position in the S/V scale is skeletal muscle which possesses nervous and humoral mechanisms capable of rendering effective large portions of the total capillary surface when active, i.e., at those times when the conditions for gas bubble formation are apt to approach the critical. Highest in the scale are cardiac muscle, kidney and liver which appear to possess a large effective capillary surface for gaseous exchanges that is available at all times.

PATHOLOGIC CONSEQUENCES OF DECOMPRESSION SICKNESS. *Central Nervous System. Pressure.* Little can be added to the classical description by Heller, Mager, and von Schrötter (80) of the neurological aspects of the syndrome of decompression sickness. The substance of their detailed description follows: After a latent period, symptoms of varying degrees of severity may appear. These may be general in scope (syncope, dizziness, aphasia), may include large body masses (paraplegia and paralysis of both lower legs, of both arms, of one arm and one leg, or both legs and one arm with a persistent spastic paralysis, loss of urinary and intestinal sphincter control), may include sensory paresthesias and hypo- or analgesia separately, or associated with motor disturbances, or may be confined to small motor or sensory units (paralysis of the left lateral rectus muscle, a single orbicular muscle, or the hypoglossal muscles on one side, labyrinthine deafness, or atrophy of the left optic papilla). Neural damage due to aeroembolism thus varies greatly, is non-specific, and unpredictable: 1, it may be extensive, covering large regions, or confined to minute structures; 2, it may be largely motor with a sensory component, or the latter may be dissociated; 3, it may be confined to areas supplied by cranial nerves or somatic nerves.

The pathologic changes in the central nervous system following damage by aeroembolism have been studied extensively in man and other animals ever since 1870. Because of numerous precautions taken to reduce the incidence of decompression sickness and because of the success of treatment when such accidents occur, the most satisfactory pathologic studies date back about half a century. The most prominent feature is the occurrence of areas of softening in the spinal cord. Although they may extend over the greater part of the spinal cord (18), they are most commonly limited to the thoracic cord (19, 29). They are less frequent in the upper lumbar segments (80, 184) and are rather infrequent in the cervical region (21, 80, 166, 184). The areas of softening are usually small and are more circumscribed in the white than in the gray matter. The small necrotic areas may fuse, and result in an apparent general softening in large regions (80). Necrosis is more common in the white than in the gray columns (80), and in the former, is most frequent in the anterior columns (21). The ascending degeneration of nerve tracts anterior to the site of multiple lesions, and the descending degeneration posterior to the region have been described by numerous pathologists (19, 80, 143, 145).

Heller, Mager and von Schrötter have described the individual lesions most

completely. In the necrotic regions of the white matter, the nerve fibers swell, degenerate and are completely resorbed. The glial meshes are spread during the edema which takes place Sclerotic areas appear, radiating into the surrounding, less damaged nerve fibers. The cavity enclosed contains cellular detritus and myelin droplets, and some phagocytic cells. The cavity then becomes more regularly outlined and the walls of the surrounding blood vessels become thicker. In the gray matter, the same processes take place, but in the surrounding regions, nerve cells in all stages of chromatolysis may occur. The necrotic cavity is not as sharply outlined in the gray as in the white matter.

While hemorrhage may occur, it is infrequent and is definitely less prominent than the regions of necrosis described above (18, 19, 80, 171, 184). It may, however, be of greater significance when it occurs in the brain (18, 29) or in the inner ear (80).

Gas bubbles have been described in pial vessels of the spinal cord, medulla and cerebrum in autopsied men, and in other animals (18, 21, 29, 80, 122, 184). As it is difficult to ascertain whether these bubbles were present where described at the time of death, the observations of Pudenz (132) and of Wagner (172) are of greater significance. The former observed in a monkey the appearance of gas bubbles first in pial arteries and then in veins in cinematographs of the cortex exposed to view by the use of a lucite calvarium. The latter made essentially the same observations in a series of cats whose pial vessels were viewed continuously before, during, and after decompression through a Forbes window. A gas bubble was observed also in the ophthalmic artery of a dog (29). Also more reliable than the earlier accounts are the preliminary observations of Gersh, who preserved the sites of gas bubbles by the freezing-drying method. Gas bubbles were found in the spinal cord of guinea pigs directly after massive aeroembolism *only* in the blood vessels; they were present also in intraneural blood vessels of the sciatic nerve. There was no evidence to support the assertion of Boycott et al. (21) that gas bubbles occur extravascularly in the central nervous system. Finally, minute gas bubbles confined entirely to the myelin sheath were present in peripheral nerves of guinea pigs decompressed from high pressure atmospheres of argon-oxygen, air, helium-oxygen, and oxygen (53).

The general pattern of neurologic symptoms, pathologic changes and the distribution of gas bubbles in the central nervous system following decompression was outlined by Bert (18). It was confirmed and extended by the magnificent work of Heller, Mager and v. Schrötter (80) which the authors have leaned on heavily, and by later workers. Gas bubbles form in the circulating blood after a short latent period following decompression. Most bubbles are filtered out by the lungs; some bubbles, however, pass through the lungs, either by means of small arterio-venous anastomoses which may be assumed to exist, or through the capillary "lakes" of Sjöstrand (147). These bubbles are small, about 25 μ in diameter. When they pass through the heart and reach the central nervous system, they occlude small arterioles of the same order of magnitude (21, 52, 132, 172). Since the circulation in the central nervous system is largely terminal in nature, the venous blood flow and blood pressure in the region is reduced, and

gas bubbles appear secondarily in the veins. Another consequence of the terminal nature of the circulation is that nerve cells and fibers in the affected region are suddenly deprived of their major source of nutrients and oxygen. They rapidly degenerate in small, focal regions corresponding to the arterioles occluded by gas bubbles. The necrotic regions may be separate, or may fuse secondarily with adjacent foci to form larger areas of softening. The extreme rapidity with which nerve cells are irreversibly injured following occlusion of the circulation has been conclusively demonstrated by numerous workers (58, 92, 93, 94, 176, 180). The unpredictability of the site and extent of the lesion, as well as the magnitude of the symptomatology, are thus related to the site of lodgement of the gas emboli, which may be isolated or adjacent to others, and to the numbers of blood vessels occluded. It cannot be stressed too strongly that there is a large element of chance as to whether a given number of gas bubbles causing neural damage will result in minor or severe symptoms, depending on the nerve cells and fibers affected. The circulation in the gray matter of the spinal cord is somewhat anastomotic, and this may account for the fact that necrotic foci occur less frequently in this portion of the spinal cord than in the white matter. Another circumstance which protects the gray matter is the fact that the capillary density, and presumably the blood flow, is greater than that in the white matter (35).

Hemorrhage is a relatively unimportant factor in neural damage of decompression sickness. Although v. Leyden believed that expanding extravascular gas bubbles induced hemorrhage by laceration of blood vessels (171), the same result may take place due to the occlusion and weakening of small blood vessels. It is also possible that vascular congestion somewhat remote from the gas bubble and the cessation of the blood flow may result in hemorrhage. This was actually observed directly in mesenteric vessels of animals injected with small gas emboli (30, 162).

The appearance of minute gas bubbles in the myelin sheath of peripheral nerves after decompression is attributable to the large volume of gas in solution in the lipid phase of the myelin sheaths. While the lipid content of the brain is only 5–8 per cent (156), certain parts may be expected to be richer in lipid. It is on this account that the earlier English investigators attributed the alleged occurrence of extravascular gas bubbles to the release of the excess lipid-soluble gas (21). Apart from the necessity for confirmation, the conclusions may be modified as a result of the later work by Campbell and Hill (23), who showed that the brain of pressurized animals contains far less gas than expected. This discrepancy in gas uptake may possibly be explained by further knowledge of blood flow and capillary density in the brain. Similar investigations may aid in interpreting several other problems without a satisfactory solution: the predominance of lesions of the spinal cord as compared with the brain, the marked predilection for injury of the thoracic cord, and the relative severity of damage to the motor systems as compared with the sensory.

Altitude. Symptoms of damage to the central nervous system following decompression to altitude are rather uncommon, and pathologic changes are even

rarer. There is no evidence that lasting changes occur in the cerebral cortex (45), in spite of the appearance of irregular EEG waves with a low frequency and high voltage during or after decompression to extreme heights (182). No residual EEG changes were detected in individuals decompressed to 35,000 feet even while they were suffering from the chokes or the bends. In subjects free from the effects of *generalized* shock, some temporary damage to the higher centers may occur, for there may be temporary syncope, scotoma, headache (45), amnesia, and auditory agnosia (106), motor aphasia, slight facial paralysis, a positive Babinski in one leg, agraphia, and a stuporous condition (22, 59, 136). Other temporary effects of decompression to altitude on the central nervous system are changes in muscle tone and in the strength of a variety of reflexes, and localized sensory losses. The incidence of *temporary* dysfunction of the central nervous system was given as 0.28 per cent, of a large series of exposures at 34,000 feet with oxygen (22). Temporary damage of the central nervous system may be even rarer than indicated, because of the difficulty of ruling out simulated effects caused basically by emotional disturbances. The only recorded instance of uncomplicated *permanent* damage to the central nervous system was reported by Smith (149) who observed paralysis in the hind limb of a dog soon after explosive decompression to 45,000 feet.

Briefly summarized, it appears that: 1. Neurologic signs of damage to the central nervous system following decompression to altitude are rare, when compared with the consequences of aeroembolism from deep diving. This relative infrequency of lesions in the central nervous system must be attributed to the relatively smaller amount of gas in solution in the body at altitude. 2. With the single exception noted above, neural damage is temporary. This may be attributed to the reduced volume of gas present in the body at altitude and to the differential between gas bubbles assumed to exist in relation to the affected regions and the high oxygen content of inspired gases. This procedure of reducing the effects of aeroembolism has been recommended for caisson workers ever since the time of Paul Bert, and is in common practice at the present time (11). 3. The neural symptoms are non-specific. As in the case of diving, this property may be attributed to the factor of chance of where a given gas-bubble freely circulating in the blood will finally lodge, occlude the local circulation and cause temporary dysfunction.

Further complicating the picture is the possibility that gas bubbles may lodge in intraneural blood vessels, and thus cause local neural (at least temporary) damage. Gas bubbles were observed in such sites in a study of acute, fatal aeroembolism in rabbits (51). An exhaustive search for gas bubbles in rabbits decompressed to altitude was made by Trowell (160). His failure to find these may be due to the inadequacy of the method of fixation employed.

Lung. Pressure. As a consequence of aeroembolism gas bubbles may occur in the pulmonary artery and its branches and in the chambers of the right side of the heart in sufficient numbers to cause death through occlusion of the circulation and asphyxia. In order to achieve asphyxia, it is necessary for enough blood vessels to be occluded to reduce the pulmonary arterial blood flow by 52–66 per

cent (64). In addition there may be pulmonary hemorrhage, edema, emphysema, and atelectasis (18, 21, 46, 80, 82). Except for minor details, the mechanisms of injury were developed by the early workers in the field. Gas bubbles arising in venous channels pass through the right auricle and ventricle to the pulmonary artery. The emboli occlude larger and smaller branches, but, contrary to earlier assertions by others, never reach the capillary bed; they are sometimes present in bronchiolar vessels (51, 56). The smallest vessels occluded were 40 μ in diameter. This corresponds with the direct observations of Chase (30) on the mesenteric vessels of living animals that gas bubbles introduced into the arterial circulation never penetrate beyond arterioles of a similar diameter. The occlusion of arterioles of similar dimensions was reported in other sites also by Curtillet (36) and by Tureen and Devine (162). The occluded blood vessels may be markedly stretched at the site of the gas bubble, as well as centrally. It should be noted in passing that the lung differs from all other organs in that gas bubbles in this structure represent emboli, with no significant contribution of local origin. While nearly all gas emboli are retained in the lung, it has been believed from the earliest writers (18) that some pass through the organ to be distributed as arterial emboli by routes which are largely hypothetical: 1, intrapulmonic arterio-venous anastomoses (36), and 2, the blood "lakes" described by Sjöstrand (147).

The relation of experimentally induced gas emboli in animals to aeroembolism appears to have been most extensively investigated by Heller, Mager and v. Schrötter (80). They found that the distribution of gas bubbles in both instances was strikingly similar. In the case of air introduced intravenously, the final distribution depended on the volume of gas injected, the rate of injection and the size of the gas bubbles.

Pulmonary hemorrhage, edema and emphysema have been shown to follow pulmonary embolization by solids (41, 165). Hemorrhage and emphysema have been attributed also to the violent rupture of alveolar walls (82). Pulmonary atelectasis as a result of decompression was first described by Bert, who attributed it to possible rupture of a vesicle at the time of decompression, thus permitting access of air under pressure to the pleural space. Other possible mechanisms are: 1, overinflation of the lung while inflating the Eustachian tube during compression; 2, expansion of air in alveoli and rupture of their walls, as a consequence of holding the breath (26); 3, escape of gas under pressure from alveoli to interstitial tissue, along the vascular sheaths to the hilus of the lung, pneumo-mediastinum, and pneumothorax (107, 108); and 4, rupture of a few alveoli into the retropleural connective tissue in the manner of a one-way valve, escape of this gas retropleurally to the hilus, pneumomediastinum, and pneumothorax (55). The mechanisms enumerated are applicable not only to the rare instances observed after decompression (4, 130, 131), but also to other equally rare situations: spontaneously (67, 120), in severe straining activity (111) and in cases of obstructed air passages, such as accompany diphtheritic laryngitis (39), following pertussis or other infections (31, 96, 111).

Altitude. Hoppe-Seyler in 1857 (84) attributed death to occlusion of pulmo-

nary capillaries by gas bubbles. The few other observations on this phenomenon were contradictory. They showed that rather few bubbles appeared in branches of the pulmonary artery, while only one bronchiolar vessel was found to be filled with gas in the series of six rabbits studied. The small number of pulmonary bubbles as compared with those in animals decompressed from pressure is explainable by the smaller quantity of gas in the body of animals decompressed to altitude. The same observation has been cited also as one of several reasons for the belief that animals decompressed to altitude tend to die of embolism of the central nervous system, rather than of pulmonary asphyxia, a condition opposite to that believed to obtain in animals decompressed from high pressure atmospheres (27, 56).

Pulmonary hemorrhage occurred irregularly in rabbits decompressed to altitude (27). Smith (149), and Berg, Baumberger, et al. (16), found that it follows more regularly after explosive decompression of experimental animals, especially, as the latter found, when decompression took place during inspiration or the last two-thirds of expiration. Only two instances of mediastinal emphysema have been reported in men after explosive decompression (112). Pneumothorax, retroperitoneal emphysema, and pulmonary atelectasis also occurred in rabbits, especially after slow decompressions. The administration of nembutal or of carbon-dioxide mixtures predisposed to pneumothorax, while preoxygenation or the administration of ammonium chloride, lactic acid, or sodium bicarbonate did not affect the incidence of atelectasis (28). The mechanism has not been investigated.

The observations on rabbits confirmed in part earlier reports by a group of British investigators who described the occurrence of atelectasis, vascular congestion, and pneumothorax in rabbits, rats, and guinea pigs decompressed to altitude (49, 50, 77, 78, 79, 159). They showed that these changes were related to the final altitude, and were unaffected by the degree of anoxia or by the CO_2 content of the blood. In unpublished experiments, the writers found that atelectasis was more common in young than in older animals.

Blood vessels. Ever since 1857 (Hoppe-Seyler (84)) it has been known that bubbles are more commonly (sometimes exclusively) present in the large veins following decompression from high pressure atmospheres, or to altitude. This was attributed by Bert to the fact that venous pressure is lower than that in arteries. Harvey et al. (72) added another factor, namely, the higher CO_2 tension of venous blood, facilitating the growth of gas bubbles. Arterial gas bubbles, when present, were attributed by Bert to passage of some gas bubbles through the lung into the left side of the heart and their distribution as emboli. Under extreme conditions, where the effect of blood pressure becomes overshadowed by other factors, one would expect gas bubbles to arise in arteries in the same manner as in veins. The distribution of gas bubbles is in fact similar to that which has been described repeatedly after the rapid intravenous injection of large volumes of air. Smaller volumes of air (or air injected at a slow rate) produce bubbles that are confined largely or almost entirely to the venous system. After decompression from high pressure atmospheres, gas is sometimes present

in the capillary bed of several tissues (52), but this seldom occurs, and then in fat tissue only, after decompression to altitude (51). The difference must be attributed to the vastly greater amount of gas in the tissues in the first condition.

Recently it has been noted by direct observations of living animals that an arteriolar and capillary constriction may occur at altitude. This has been observed to take place in man, rats and frogs (99, 104, 121), though not in cats (125). Reed and Blinks (138) found that vasoconstriction was not related to bubble formation. They also found that it could be abolished by nerve section and by exercise. It was correlated in man with the prolonged refilling time of the finger nail bed and of the arm vein (99), and with reduced skin temperature (157). This group found evidence that the latter phenomenon is related to susceptibility to bends.

Knisely (99) failed to see evidence of intravascular sludge in men exposed to altitude, although he did describe vasoconstriction. On the other hand Patek, who failed to note any change in the caliber of blood vessels in cats, observed sludging (125). This may have been due to the anoxic condition of the animals.

Sludging was observed directly in the pial vessels of living animals injected intravenously with air or decompressed from high pressure atmospheres. It occurred after vascular occlusion by gas bubbles, presumably as a consequence of reduced circulation and anoxia (170, 172). End (44) regarded agglutination of red blood corpuscles as the causative factor of decompression sickness, and gas bubbles as a secondary event. Swindle et al. (154) attributed disturbed function of decompression sickness to the occurrence of fragile, non-gaseous plasma flocculates. Jacobs and Stewart (88) investigated these possibilities by studying rat blood exposed to high pressure atmospheres and then decompressed. They found no clear evidence of any change in sedimentation rate, tendency to rouleaux formation or in the aggregates of rouleaux. They failed to observe true agglutination of red blood cells, but did describe a tendency for blood platelets to form aggregates about small gas bubbles and to appear free in this form later. Such platelet aggregates could conceivably lead to occlusion of small vessels in certain regions of the body. Gersh and his co-workers paid particular attention to the possible occurrence of all of these phenomena and failed to detect them in animals decompressed from high pressure atmospheres or to altitude. As the method employed (fixation by freezing and drying) could be confidently expected to preserve such flocculates and aggregates, the failure to find them returns the causation of decompression sickness to the conventional realm of gas bubbles.

Fat Tissue. Pressure. Experience in diving or caisson operations is that accidents occur more commonly in fat workers than in lean. Heller, Mager and v. Schrötter (80) reported that animals rich in fat in the subcutaneous tissues, mesentery and pericardium are more prone to show abnormalities than lean animals. This was tested for the first time by Boycott and Damant (20), when they found that the time of survival of rats and guinea pigs after decompression from a stated pressure was related to their total fat content. This was confirmed in another way by observing the highest pressure which guinea pigs could survive when decompressed rapidly, and relating it to their total fat content. For ex-

ample, fat guinea pigs (average fat content 31.9 per cent) survive decompression from a pressure of 45 lbs/sq. in., but die when brought up from 60 lbs/sq. in.; on the other hand, lean guinea pigs (average fat content 8.2 per cent) usually survive decompression from an atmosphere at 90 lbs/sq. in. but die when brought up from 105 lbs/sq. in. (52). Using another criterion (i.e., the time of appearance of the first gas bubble), the tendency for bubble formation in resting cats was also found to be related to the fat content of the animal when it was decompressed from moderately increased pressures. But in active animals, or at greater pressures, this relationship was masked by other factors (70).

Fat tissue containing numerous gas bubbles appears foamy, like whisked white of egg (82). The bubbles may be intravascular as well as extravascular (20). As a consequence, in surviving animals, large masses of necrotic fat may be present, especially below the kidneys. In late stages, the necrotic regions are surrounded by a zone of giant-cells, with some of the fat converted to calcium soap (21). In guinea pigs, minute gas bubbles may occur in the intracellular fat inclusions of fat cells. In somewhat more severely affected animals, macroscopic bubbles which are extravascular and extracellular may occur. Intravascular gas bubbles may be present in fat tissue in the absence of the aforementioned bubbles. Hemorrhage and vascular distention are associated frequently with the occurrence of extravascular gas, especially in fat animals (52). Thus, it may be assumed that, when purpuric subcutaneous spots appear in divers, gas bubbles have occluded some blood vessels and vascular congestion and hemorrhage have taken place.

The genesis of gas bubbles in fat tissue may be reconstructed in some such manner as follows: Fat cells dissolve excess amounts of gas, especially nitrogen, during pressurization. During and after decompression, many fat cells enlarge due to the appearance in the fat inclusions of minute intracellular bubbles. Meanwhile dissolved gas passes from the fat inclusions to the tissue fluid and the circulating blood. Depending on a number of factors, gas continues in solution or forms visible circulating bubbles. When these exceed the diameter of the blood vessels, they occlude the circulation. Meanwhile, in some regions, the cells increase in volume more markedly due to the increased number of intracellular gas bubbles. These distended cells then rupture and discharge their contents into an irregularly outlined intercellular (extravascular) bubble which contains cellular debris, fat, and gas under pressure. About this time, some blood vessels occluded by gas bubbles are distended with blood, and others may rupture.

This hypothetical reconstruction of the events leading to the formation of gas bubbles in fat tissue is supported by evidence along at least three lines that certain physical factors are involved. These are as follows: 1. The increased solubility of gases in fat as compared with tissue fluid (see p. 368). Biological evidence may be cited. The occurrence of gas bubbles in fat tissue and their number are directly related to its lipid content. Extravascular gas bubbles occur only in lipid rich structures (fat tissue, adrenal cortex, myelin sheaths of peripheral nerves) (52). Finally, gases which are especially soluble in lipid (such as argon),

are much more effective in producing gas bubbles in these structures, while gases which have a low $\frac{\text{fat}}{\text{water}}$ index of solubility are relatively ineffective (53). 2. The increased amount of gas in solution in fat is directly related to the number of gas bubbles formed in fat tissue. This was illustrated by the greater severity of the effects of decompression when animals are pressurized at higher pressures. It was also shown by a comparison of the effects of argon-oxygen, and nitrogen-oxygen mixtures. Although the relative solubilities in lipid and tissue fluid of argon and nitrogen are nearly identical, the amount of the former in solution is over twice as great. The number of gas bubbles appearing in fat tissue is far greater when argon is used than when nitrogen is used. In both examples, variations in the number of gas bubbles occur although the size of the gas bubbles remains the same. The size of the gas bubbles is probably referable to some tissue property such as elasticity. 3. The poor capillary density of fat tissue is an important factor predisposing to bubble formation, since excess gas cannot be removed rapidly enough to avoid high ΔP values in the tissue (see p. 362).

Altitude. In contrast with the serious consequences of excess body fat for divers, fat is not of great significance in bubble formation on decompression to altitude. Harvey et al. (70) and Catchpole and Gersh (27) failed to find any correlation between the fat content of animals decompressed to altitude and the time of first appearance of intravascular gas bubbles, the severity of the symptoms, or the time of survival.

Extravascular gas bubbles do not occur. Gas bubbles have been observed in veins draining fat depots but less commonly than in blood vessels draining muscles (70). They were found to be somewhat more numerous in fat tissue than in others by Gersh and Catchpole (51). Fat is the only tissue where gas bubbles were present in the capillary bed.

Muscle. Pressure. Little emphasis is placed in the literature on muscular lesions following decompression except as a consequence of neurological damage, or in the case of regional infarct (169). Only in recent years has muscle acquired great significance in theories on the origin of gas bubbles. Bert (18) found bubbles to be present only in intermuscular fascia, and Boycott et al. (21) were unable to see gas bubbles in skeletal muscle. Harvey et al. (70) found that in anesthetized cats bubbles do not appear in veins draining largely muscular regions, although they are present in vessels from fat depots. On the other hand, in stimulated animals, bubbles arise only from large muscle masses. Both the Princeton and the California groups agree that violent activity increases the number of bubbles (70, 177). In normal, decompressed guinea pigs, gas bubbles were present in muscle in arteries, capillarics, and veins of all dimensions, but they were fewer in number than in blood vessels of fat depots (52). Under the same conditions, an equally small number of blood vessels were occluded after decompression from a helium-oxygen mixture, a larger number were observed after argon, and none were observed after oxygen.

Muscle is regarded as an important site of bubble formation, the gas phase appearing in regions of mechanical tension, whether this is induced by passive

movement, stretching, injury or stimulation. This is attributed to the decreased P obtaining near tendon attachments due to the local stresses (68, 110). It should be stated that in guinea pigs, although blood vessels near the knee joint were not noted for their density of gas bubbles, those accompanying the *long* tendons near the ankle joint displayed gas bubbles prominently (56). Also important are the increased volumes of gas in solution as a result of compression. The increased gas content appears to overshadow the facilitating action of carbon dioxide.

Altitude. In anesthetized, decompressed animals, relatively few gas bubbles appear, and these seem to originate in fat depots. In stimulated limbs, however, the numerous bubbles appear to arise primarily from muscle (27, 28, 69, 110). By direct visualization, no bubbles were found in muscles of decompressed rats, although they were present at autopsy in the vena cava and other large veins (104). Arterioles were observed to be constricted. On the other hand, small numbers of bubbles were described in decompressed rabbits in blood vessels of larger-than-capillary dimensions. The factors responsible for bubble formation are primarily: 1, local production of CO_2 in regions of muscular activity, so that gas nuclei or minute bubbles in regions of high CO_2-tension tend to expand rapidly; 2, mechanical tension induced by passive movement, stretching, injury or electrical stimulation, which results in regions of low pressure in the neighborhood of insertions (70, 72, 110, 177, 179).

Bone Marrow, Bone, Periosteum, Joints. Bone marrow. The occurrence of gas bubbles in bone marrow has been described only in acute preparations of guinea pigs decompressed from compressed air (56). It was rather unexpected to find that in yellow marrow which was equivalent to lipid-rich fat, intravascular bubbles were exceedingly rare and extravascular bubbles were absent while the fat of the same animals was filled with numerous bubbles in both sites. Only under more severe conditions of pressurization did extravascular gas bubbles appear, together with a larger number in arteries, sinusoids, and veins. The extravascular gas bubbles were frequently larger than in fat, more irregular, and with a tendency to be concentrated near bony spicules. Sometimes, they seemed to dissect the endosteum from the marrow tissue. Arterial occlusion by gas is probably the result of embolism. Venous and sinusoidal gas bubbles probably arise locally in the bone marrow due to the slow circulation (24), the increased amount of nitrogen stored in the inclusions of the fat cells, and the low blood pressure in the blood channels. The extravascular distorting bubbles probably arise in a manner similar to that described for fat tissue. Both the large size of the extravascular bubbles and their irregular distribution may be due to the intermittent flow of blood through some sinusoids. This would result in a high uptake of excess gas in some regions and a large t locally, with bubble formation, and in other regions, low uptake of gas and no bubble formation. The tendency for more bubbles to occur in relation to bony spicules and to dissect the endosteum may be related to the finding that some crystals tend to favor the origin of gas bubbles from gas nuclei (126). It should be emphasized again that gas bubbles in bone marrow are relatively scarce, and then even so they

occur only in extreme conditions. This is probably due to the retarded rate of saturation of bone marrow with excess gas and the reduced circulation.

Bone and periosteum. Pressure. Bone necrosis as a result of decompression sickness was first noted in 1888 (163). The literature on the subject has been thoroughly reviewed by Kahlstrom, Burton and Phemister (95). Additional information has been supplied by Phemister (127), Coley and Moore (32), and Rendich and Harrington (139). The most detailed microscopic study was made by Kahlstrom et al. They found that when necrotic bone was situated in the epiphysis and bordered on joints, there were varying degrees of collapse of the weight-bearing portions, and invasion of the necrotic regions with new bone or calcification. The overlying articular cartilaginous surfaces were replaced by fibro-cartilage, resulting in arthritis deformans. When the necrotic bone was located in the diaphysis, or in the epiphysis removed from articular surfaces, replacement by new bone or calcification also took place, but collapse and deformation did not occur. Since these observations were made several years after the original trauma, the mechanism could not be determined exactly. They postulated two general methods by which gas bubbles could cause bone atrophy: 1, gas bubbles form in bone marrow and interrupt the circulation of bone, and 2, gas emboli in numbers of smaller branches of the nutrient artery occlude the osseous circulation. Both mechanisms were demonstrated in work on decompressed guinea pigs. In some instances, gas-filled Haversian vessels were observed (56). It should be pointed out that bone changes in caisson workers and divers are relatively uncommon, probably for the same reasons that were given to account for the scarcity of gas bubbles in bone marrow. Gas bubbles are even scarcer in the periosteum, where, only rarely, gas distended small blood vessels were seen (56).

Altitude. Only one systematic effort has been made to detect bone damage in men decompressed to altitude. No radiographic evidence of such effects was found in a rather large number of subjects with 5–100 hours' exposure at 35,000 feet or higher. Follow-up x-rays several years later were recommended (135). The probability is that bone damage will be exceptional, for the same reasons that neural damage in aviators is scarcer than in divers (see p. 375). Another solitary report on the appearance of periosteum in living animals failed to disclose any evidence of bubbles in this site (125). However, the altitude at which the animals were studied, and the rate of decompression may have been unfavorable.

Joints and periarticular structures. Pressure. Experimental studies utilized morphological methods for the demonstration of gas bubbles or their effects, and x-rays to locate gas bubbles photographically. In guinea pigs decompressed from high pressure atmospheres, gas bubbles may be present in the joint fat. But even when they are absent, blood vessels in the region, especially those adjacent to the long tendons are frequently occluded and distended by gas bubbles (56). With x-rays, only the larger blood vessels which were completely filled with gas, were recognizable. Bubbles in fat or connective tissue were indistinguishable, or when they were very numerous, were recognizable almost always as a general x-ray shadow. In only one instance, in which shadows suggestive of

bubbles were found, only a minute fraction of all the bubbles were recognizable as such, because of the small size of the bubbles, poor resolution and overlay (57).

Altitude. The only morphological studies reported were on the joints of living animals. The failure to find gas bubbles in the periarticular tissues may have been due, however, to experimental conditions (125). Webb and co-workers (174) attempted to correlate gas in tissues of the knee with pain, and found that pain is unrelated to gas in the joint space or in the suprapatellar bursa. Better statistical relationships of pain were found with x-ray photographs correlating, in decreasing order of success, with 1, streaks in tissues posterior to the joint capsule; 2, gas bubbles immediately posterior to the joint space, and 3, gas bubbles in the infrapatellar space. The streaks appeared to be extravascular collections of gas in the intermuscular fascial planes of the posterior tissues of the leg. The gas seemed to be disposed as a collection of minute bubbles and to have no relation to the *larger* blood vessels or nerves of the region. Similar findings were reported by Thomas and Williams (155), who pointed out that "streaking" may occur in the absence of pain, and that there was no characteristic x-ray picture with which pain is correlated. Lund et al. (105, 106) came to the same conclusion in their study on the origin of pain in bends, and expressed the belief that the demonstration of gas by radiographic measures alone can be expected to help very little in supporting theories on the intimate origin of pain. A similar attitude was expressed by Gersh (57), who pointed out the difficulties in interpreting x-ray photographs of gas bubbles in terms of finer anatomical structures.

Adrenal Gland. Pressure. Gas bubbles have been detected in all layers of the adrenal cortex, as well as in the medulla. In addition, they have been observed in the capsular arteries, in the medullary veins and its tributaries, and occasionally in the reticular sinusoids. In all layers of the cortex, but most frequently in the fascicular zone, gas bubbles are present also extravascularly. Three major pathological consequences are 1, the occurrence of hemorrhages in the vicinity of the gas bubbles; 2, rupture of cortical cells in the region, and 3, vascular congestion in the sinusoids peripheral to the gas bubbles, with cessation or reduction of blood flow (52).

Two main factors are important in the origin of gas bubbles in the adrenal gland: the high content of lipid and certain peculiarities of the vascular pattern. The following sequence of events may be reconstructed: Gas bubbles begin as small occluding emboli in capsular arterioles and as clusters in the sinusoids of the reticular zone after the blood has passed through the fat- (and nitrogen) rich fascicular zone. These may grow and further occlude the circulation in the cortex. These areas of reduced cortical blood flow are small, and surrounding regions may have a normal circulation. With the two regions having different gas tensions, the gas bubbles continue to grow in the former region, and destroy the cortical organization, particularly in the fascicular zone, where the fat and dissolved gas is richest. The general disruption of structure is accompanied by hemorrhage, and also by a further sinusoidal congestion. Meanwhile, in the adjacent unaffected region, gas in solution is removed by the circulation either in solution or as minute bubbles which do not clog the sinusoids. Gas bubbles

arising in this manner may, in the capacious venous channels of the medulla, where the blood pressure may be smaller, expand to form large intravascular gas bubbles which may seriously interfere with the circulation throughout the gland.

Although gas bubbles appear in the adrenal gland of decompressed guinea pigs, it should be emphasized that they form somewhat less readily than in fat tissue. It is difficult to evaluate the clinical significance of the pathological findings. The temporary withdrawal of functioning cortical tissue may be related to the sense of exhaustion which divers frequently report after immersion.

Altitude. Gas bubbles were observed only in capsular arterioles and in some of the large venous medullary channels in rabbits; none were found in sinusoids or extravascularly (51).

Kidney. Pressure. Destruction of renal tubules of the cat was prominently figured by Boycott, Damant and Haldane (21). The proximal convoluted tubules of the cat are virtually unique because they are exceedingly rich in lipid, and it is unlikely that similar effects will occur in the kidney of other animals. In support of this assertion is the rapidity of the blood flow, and the brief desaturation time of renal tissue (see p. 370).

Altitude. Gas bubbles were present only in renal blood vessels of large caliber in rabbits decompressed rapidly (51).

Liver. Pressure. Boycott, Damant and Haldane (21) found numerous bubbles issuing as a froth from the cut surface of the liver of decompressed animals. These were believed to be confined to blood vessels. With methods less subject to artefact, very few microscopic gas bubbles were seen infrequently and only in central veins (52). Gas bubbles were common after decompression from a high pressure atmosphere of argon-oxygen (when some bubbles were observed in sinusoids), were rare after helium-oxygen mixtures, and were absent after oxygen (53).

A peculiar finding was the appearance of watery vacuoles in the cytoplasm of liver cells. The spherical vacuoles displace and distort the nucleus, and contain no fat stainable with Sudan III. They were present in liver cells not only after decompression from compressed air, but also from atmospheres of argon-oxygen, helium-oxygen, and oxygen. Their significance is unknown (52, 53).

The scarcity of bubbles in the liver may be explained by the work of Campbell and Hill (25) who found that liver has a rather slow half-saturation time. This may be due to the fact that most of the hepatic blood has already passed through a capillary bed and given up most of the excess gas. In addition to the reduction of the gas volume in the liver due to this factor, the nature of the arterial supply may contribute to the total picture during and after decompression. The less saturated arterial blood emptying into the sinusoids would tend to reduce the gas tension of blood in them and in the central vein and thus reduce the tendency for bubble formation.

Altitude. In decompressed rabbits, microscopic bubbles were observed very seldom, in the central vein and in the branches of the portal vein. The intracellular vacuoles observed in animals decompressed from high pressure atmospheres did not appear. They have, however, been described in men who died of the effects of exposure to altitude and in animals subjected to sudden reduction in

atmospheric pressure (103, 158). Ladewig gave good evidence to indicate that they may be related to low oxygen tension and anoxia rather than the effects of decompression. On the other hand, it is difficult to see how anoxia would play a rôle in the appearance of the vacuoles of guinea pigs decompressed rapidly enough from high pressure atmospheres to result in death in one to three minutes.

Spleen. Pressure. Gas bubbles occur in splenic sinusoids, arteries, and veins, in decreasing order of frequency, of guinea pigs decompressed rapidly from high air pressures (56). The sinusoidal bubbles are probably related to the intermittency of blood flow and possibly the reduced rate of blood flow in this structure. Both factors may result locally in regions of high gas tension (in closed sinusoids or in those with sluggish circulation) adjacent to regions of low gas tension (in narrow sinusoids with rapid circulation). If the volume of gas in the adjacent tissue is large enough, the gas bubbles grow and cause large, tearing defects to appear. The arterial bubbles are probably the result of gas emboli. The venous bubbles probably originate in situ because of the increased gas tension and decreased blood pressure.

Altitude. Fewer gas bubbles were present in the sinusoids of rabbits decompressed to altitude; they were present also in arteries and veins. In two rabbits, large, subcapsular, tearing bubbles were visible (51). The factors responsible for their appearance are probably the same as those described above.

Site of Origin of the Bends and the Chokes. Although much has been written on the subject, the literature dealing with unequivocal, direct evidence is distressingly scanty. If one makes the logical assumption that the pain of the bends and the chokes is due to gas bubbles, then from the point of view of the pathologist, those regions which are prone to bubble formation in man or other animals under less extreme conditions of decompression acquire the greatest significance as possible sites of the origin of pain. Whether pain results from distortion of nerve endings or nerve fibers by intra- or extravascular gas bubbles, or from local ischemia arising from reduced or occluded circulation is unknown. Gas bubbles have been described in muscle, intermuscular fascia, fat tissue, nerves, bone marrow, periosteum, and in fascial tissues adjacent to joints or to the long tendons. Except for fat tissue (where intracellular bubbles may occur in the cytoplasmic fat inclusion and also extravascularly) and for the myelin sheath of nerve fibers, no other extravascular gas bubbles have been noted in organs of locomotion. Even in these sites, gas bubbles may be seen in blood vessels in animals decompressed from high pressure atmospheres before they appear extravascularly; in animals decompressed to altitude, only intravascular bubbles occur. In all other regions, gas bubbles are confined entirely to the blood vessels. Thus, granting the primary assumption, the conclusion seems inescapable that the site of origin of gas bubbles is primarily vascular, though it cannot be denied that the extravascular gas bubbles may in addition cause pain by distortion. The evidence is cited in previous sections for the conclusion that the blood vessels most likely to be involved in the causation of the bends are those in the vicinity of the joints or tendons near their origin or insertion. However, others located in

nerves, muscles, or fasciae (including periosteum) may also be additional sites where pain may originate. A similar analysis of the anatomic basis of the pain of the chokes leads to the conclusion that the most likely sites of origin are the branches of the pulmonary arteries.

The conclusions derived above must be modified to a certain extent in the light of evidence showing that vascular occlusion by a gas bubble is accompanied by vascular congestion (30). If pain arises from vascular distention, then the same blood vessel or branches thereof somewhat removed from the gas bubble may be stretched to nearly the same extent. If pain arises from ischemia, then any point in the anoxic region may result in stimulation of the proper nerve endings or fibers.

Effects of Exercise and Drugs on Aeroembolism. Pressure. The conflicting claims of the effects of exercise on divers may be resolved by a critical analysis of studies by Harris et al. (68) and by Harvey et al. (70). The first group found that the minimal pressure which results in the liberation of bubbles after decompression is markedly lowered as a result of exercise. It may be assumed that as the pressure is increased greatly over this minimum value, the amount of gas present in the body outweighs in importance the decreased P which accompanies muscular activity. The result of this is that it would be difficult in practice to show that there is any relationship between exercise and amount of gas bubbles formed on decompression from higher pressures.

The effects of exercise were illustrated in another way by the use of anesthetics. Anesthetized animals, probably as a result of decreased muscular tone, had a much higher minimum effective pressure. Harvey and his colleagues (70) made similar observations, and found that a clear relationship between exercise and bubble formation existed on decompression from lower pressures, but that this is masked after higher pressures. From these results it is possible to understand that in divers, where the time under pressure, the pressure, and the rate of decompression are not often duplicated in experiments, some workers may claim that exercise is harmful, ineffective, or even beneficial in preventing or ameliorating the bends. Harris et al. (68) also reported that the use of anesthetics reduced bubble formation markedly, probably as a result of decreased muscular tone.

Oxygen administration is used effectively for the prevention and treatment of the more serious effects of aeroembolism in divers (5, 11, 89). This is confirmed by experiments on guinea pigs, in which far fewer gas bubbles in all sites were noted after decompression from oxygen as compared with nitrogen (53). This is difficult to understand, as the solubility properties of both gases are very nearly identical. Certain phases of decreased aeroembolism may be explained by assuming that oxygen of gas emboli is used metabolically, thus, in effect, reducing t at the site of bubble formation.

Altitude. The accelerating effects of exercise on bubble formation in animals at altitude have already been described. They correspond with the well-known increased severity of symptoms induced in men decompressed to altitude under controlled conditions. The increased tendency to bubble formation is attributed

to decreased P in the region of tendon attachments, and increased local t values due to excessive accumulation of CO_2

Preoxygenation is effective in delaying the onset of symptoms of aeroembolism in men, and in protecting animals against bubble formation (see p. 369). This effect is achieved presumably by substituting a metabolizable gas (reducing t) for nitrogen, which is largely eliminated during the period of preoxygenation.

The administration of ammonium chloride and lactic acid was also found to reduce the tendency to bubble formation in decompressed rabbits (28). In man (2, 3, 113) ammonium chloride increases tolerance to anoxia and improves subjective symptoms. The underlying mechanism in the latter studies appear to be a shift in the blood pH towards the acid side and correction of the alkalosis caused by hyperventilation and loss of CO_2. To the extent to which CO_2 contributes to the initiation of bubble formation at altitude, a lowering of the arterial CO_2 tension would be expected to hinder bubble formation. The administration of moderate concentrations of CO_2 or of sodium bicarbonate did not have a noticeable effect on bubble formation (28, 110, 177). Accordingly, it was not surprising that in men exposed for two hours at a simulated altitude of 38,000 feet, the symptoms were of the same order whether the subject breathed pure oxygen or oxygen diluted by CO_2 (60).

There has been no uniform interpretation of the results of administering vasodilator drugs. Aminophyllin has been claimed to alleviate, or to reduce the severity of the bends (180); on the other hand, it had no effect on bubble formation in animals (138). A favorable effect of d-amphetamine in reducing the incidence of incapacitating bends has been both proposed (87) and denied (141).

SUMMARY

The evidence is overwhelming that gas bubbles are the primary pathogenetic agent in eliciting the pathologic effects of decompression sickness. Whether they occur after decompression from high pressure atmospheres or to altitude, gas bubbles are chiefly intravascular, and they are held to be responsible for nearly all important phases of the syndrome of decompression sickness. Extravascular gas bubbles occur also under certain severe instances of decompression from high pressure atmospheres, but they are restricted to certain lipid-rich structures. The pathological effects may be vastly greater after decompression from high pressure atmospheres than to altitude. These are described in detail for the various tissues and organs of the body. An attempt has been made to relate earlier and recent findings to each other, and to the causative factors. Basic to an understanding of the mechanisms involved in the syndrome is a consideration of the physical factors responsible for the uptake and elimination of excess gas in the body. The following physical factors were subjected to analysis: intrinsic factors such as blood pressure, blood flow, tissue permeability, tissue activity, the chemical composition of tissues and the solubility of gases in body components, and extrinsic factors such as the type, rate and extent of decompression applied.

REFERENCES[1]

(1) Armstrong, H. G. Principles and practice of aviation medicine. Williams & Wilkins Co., Baltimore, 1939.
(2) Barach, A. L., W. L. Bloom, M. Eckman, C. Rule and C. C. Rumsey. The acid-base balance of the blood following ingestion of ammonium chloride. Committee on Aviation Medicine, Report No. 240, National Research Council, December 24, 1943.
(3) Barach, A. L., et al.; R. Hodes and M. G. Larrabee. Summary of reports on the effect of ammonium chloride on altitude tolerance. Committee on Aviation Medicine, Report No. 96, National Research Council, December 1942.
(4) Behnke, A. R. Analysis of accidents occurring in training with the submarine "lung". U. S. Nav. Med. Bull. 30: 177, 1932.
(5) Behnke, A. R. The application of measurements of nitrogen elimination to the problem of decompressing divers. U. S. Nav. Med. Bull. 35: 219, 1937.
(6) Behnke, A. R., Jr. Physiologic studies pertaining to deep sea diving and aviation, especially in relation to the fat content and composition of the body. The Harvey Lectures, 198, 1941–1942.
(7) Behnke, A. R. The absorption and elimination of gases of the body in relation to its fat and water content. Medicine 24: 359, 1945.
(8) Behnke, A. R. Decompression sickness incident to deep sea diving and high altitude ascent. Medicine 24: 381, 1945.
(9) Behnke, A. R. The absorption and elimination of gases of the body in relation to its fat and water content. Medicine 24: 359, 1945.
(10) Behnke, A. R., Jr., B. G. Feen and W. C. Welham. The specific gravity of healthy men. J. A. M. A. 118: 495, 1942.
(11) Behnke, A. R., L. A. Shaw, A. C. Messer, R. M. Thomson and E. P. Motley. The circulatory and respiratory disturbances of acute compressed-air illness and the administration of oxygen as a therepeutic measure. Am. J. Physiol. 114: 526, 1936.
(12) Behnke, A. R., R. M. Thomson and L. A. Shaw. The rate of elimination of dissolved nitrogen in man in relation to the fat and water content of the body. Am. J. Physiol. 114: 137, 1935.
(13) Behnke, A. R. and T. L. Willmon. Gaseous nitrogen and helium elimination from the body during rest and exercise. Am. J. Physiol. 131: 619, 1940.
(14) Behnke, A. R. and T. L. Willmon. Cutaneous diffusion of helium in relation to peripheral blood flow and the absorption of atmospheric nitrogen through the skin. Am. J. Physiol. 131: 627, 1940.
(15) Behnke, A. R. and O. D. Yarbrough. Physiologic studies of helium. U. S. Nav. Med. Bull. 36: 542, 1938.
(16) Berg, W. E., J. P. Baumberger, F. Crescitelli, S. Rapaport and P. O. Greeley. Explosive decompression: lung damage correlated with the respiratory cycle in explosive decompression. Committee on Aviation Medicine, Report no. 173, National Research Council, August 16, 1943.
(17) Berg, W. E., M. Harris, D. M. Whitaker and V. C. Twitty. Additional mechanisms for the origin of bubbles in animals decompressed to simulated altitudes. J. Gen. Physiol. 28: 253, 1945.
(18) Bert, P. La pression barométrique; recherches de physiologie experimentale. Paris, G. Masson, 1878.—Barometric pressure. Transl. by M. A. Hitchcock and F. A. Hitchcock, Columbus, Ohio, College Book Co., 1943.
(19) Blanchard and Regnard. Cited by Heller, Mager and v. Schrötter, 1881.
(20) Boycott, A. E. and G. C. C. Damant. Experiments on the influence of fatness on susceptibility to caisson disease. J. Hyg. 8: 445, 1908.
(21) Boycott, A. E., G. C. C. Damant and J. S. Haldane. The prevention of compressed-air illness. J. Hyg. 8: 342, 1908.

[1] To June, 1946.

(22) Brown, G. A., C. H. Cronick, H. L. Motley, E. J. Kocour and W. O. Klingman. Nervous system dysfunction in adaptation to high altitude and postflight reactions. War Med. **7**: 157, 1945.
(23) Campbell, A. and L. Hill. Concerning the amount of gas in the tissues and its removal by breathing almost pure oxygen. J. Physiol. **71**: 309, 1931.
(24) Campbell, A. and L. Hill. Studies in saturation of the tissues with gaseous nitrogen. I. Rate of saturation of goats' bone-marrow in vivo with nitrogen during exposure to increased atmospheric pressure. Quart. J. Exper. Physiol. **23**: 197, 1933.
(25) Campbell, J. A. and L. Hill. Studies in saturation of tissues with gaseous nitrogen. III. Rôle of saturation of goat's brain, liver and bone marrow, in vivo, with excess nitrogen during exposure to +3, +4, and +5 atmospheres pressure. Quart. J. Exper. Physiol. **23**: 219, 1933.
(26) Case, E. M. and J. B. S. Haldane. Human physiology under high pressure. I. Effects of nitrogen, carbon dioxide, and cold. J. Hyg. **41**: 225, 1941.
(27) Catchpole, H. R. and I. Gersh. Physiological factors affecting the production of gas bubbles in rabbits decompressed to altitude. J. Cell. and Comp. Physiol. **27**: 15, 1946.
(28) Catchpole, H. R. and I. Gersh. Bubble formation in rabbits decompressed to altitude: effect of preoxygenation, electrical stimulation, and some pharmacological factors. J. Cell. and Comp. Physiol. **27**: 27, 1946.
(29) Catsaras. Cited by Heller, Mager and v. Schrötter, 1886.
(30) Chase, W. H. Anatomical and experimental observations on air embolism. Surg., Gynec. and Obstet. **49**: 569, 1934.
(31) Clark, E. and M. J. Synnott. Influenza-pneumonia cases showing gas in fascial tissues. Am. J. Med. Sci. **157**: 219, 1919.
(32) Coley, B. L. and M. Moore, Jr. Caisson disease with special reference to the bones and joints. Ann. Surg. **111**: 1065, 1940.
(33) Cobb, S. Cerebral circulation: a critical discussion of the symposium. Assn. Res. Nerv. and Ment. Disease **18**: 719, 1938.
(34) Cobb, S. and J. H. Talbot. Studies in cerebral circulation. II. A quantitative study of cerebral capillaries. Trans. Assn. Amer. Phys. **42**: 255, 1927.
(35) Craigie, E. H. The comparative anatomy and embryology of the capillary bed of the central nervous system. Assn. Res. Nerv. and Ment. Disease **18**: 3, 1938.
(36) Curtillet, E. L'embolie gazeuse artérielle. J. de Chirurg. **53**: 461, 1939.
(37) Daly, I. de B., P. L. Eggleton, S. R. Elsden, C. O. Hebb and O. A. Trowell. The significance of safe decompression rates for divers in relation to safe rates of ascent for airmen with special reference to the effect of explosive decompression. Flying Personnel Research Committee, Rept. no. 200, Nov. 1940.
(38) Dean, R. B. The formation of bubbles. J. Appl. Physics **15**: 446, 1944.
(39) Dolgopol, V. B. and M. E. Stern. Interstitial emphysema of the lung with spontaneous pneumothorax and subcutaneous emphysema. Arch. Otolaryngol. **31**: 140, 1940.
(40) Dumke, P. R. and C. F. Schmidt. Quantitative measurements of cerebral blood flow in the macaque monkey. Am. J. Physiol. **133**: 421, 1943.
(41) Dunn, J. S. The effects of multiple embolism of pulmonary arterioles. Quart. J. Med. **13**: 129, 1920.
(42) Duyff, J. W. and Bouman. Zellforsch. **5**: 596, 1927, cited from Krogh.
(43) Eggleton, P., S. R. Elsden, J. Fegler and C. O. Hebb. A study of the effects of rapid decompression in certain animals. J. Physiol **104**: 129, 1945.
(44) End, E. The use of new equipment and helium gas in a world record dive. J. Ind. Hyg. and Toxicol. **20**: 511, 1938.
(45) Engel, G. L., J. Romano, J. P. Webb, E. B. Ferris, Jr., H. W. Ryder and M. A. Blankenhorn. Absence of demonstrable injury to the central nervous system

after repeated experiencing of decompression sickness. Committee on Aviation Medicine. Report no. 263, National Research Council, March 1, 1944.

(46) ERDMAN, S. The acute effects of caisson disease or aeropathy. J. Med. Sci., N.S. **145:** 520, 1913.

(47) ESSEX, H. E., J. F. HERRICK, E. J. BALDES AND F. C. MANN. Influence of exercise on blood pressure, pulse rate, and coronary blood flow of the dog. Am. J. Physiol. **125:** 614, 1939.

(48) FERRIS, E. B., W. E. MOLLE AND H. W. RYDER. Nitrogen exchange in tissue components of man. Committee on Aviation Medicine, Report no. 60, National Research Council, July 15, 1942.

(49) FEGLER, J. Observations on guinea-pigs explosively decompressed to equivalent heights of 41,000 to 45,000 feet. Flying Personnel Research Committee, no. 349, August, 1941.

(50) FEGLER, J., C. HEBB AND W. MISSIURO. Postmortem conditions of the lung characteristic of deaths at high altitudes. Flying Personnel Research Committee, no. 575, February, 1944.

(51) GERSH, I. AND H. R. CATCHPOLE. Appearance and distribution of gas bubbles in rabbits decompressed to altitude. J. Cell. and Comp. Physiol. in press, 1946.

(52) GERSH, I., G. E. HAWKINSON AND E. N. RATHBUN. Tissue and vascular bubbles after decompression from high pressure atmospheres—correlation of specific gravity with morphological changes. J. Cell. and Comp. Physiol. **24:** 35, 1944.

(53) GERSH, I., G. E. HAWKINSON AND E. H. JENNEY. Comparison of vascular and extravascular bubbles following decompression from high pressure atmospheres of oxygen, helium-oxygen, argon-oxygen and air. J. Cell. and Comp. Physiol. **26:** 63, 1945.

(54) GERSH, I. AND M. A. STILL. Blood vessels in fat tissue. Relations to problems of gas exchange. J. Exper. Med. **81:** 219, 1945.

(55) GERSH, I. Pneumothorax and extrapulmonic emphysema in cats exposed to oxygen under pressure. Project X-192, Report no. 5, Naval Medical Research Institute, October 30, 1944.

(56) GERSH, I. Gas bubbles in bone and associated structures, lung and spleen of guinea pigs decompressed rapidly from high pressure atmospheres. J. Cell. and Comp. Physiol. **26:** 101, 1945.

(57) GERSH, I. Correlation of x-ray and gross observations on gas bubbles in guinea pigs decompressed from high pressure atmospheres. J. Cell. and Comp. Physiol., in press, 1946.

(58) GILDEA, E. F. AND S. COBB. The effects of anemia on the cerebral cortex of the cat. Arch. Neurol. and Psychiat. **23:** 876, 1930.

(59) GOGGIO, A. F. AND G. H. HOUCK. Physiologic abnormalities and pathologic changes following exposure to simulated high altitudes. War Med. **7:** 152, 1945.

(60) GRAY, J. S., R. L. MASLAND AND S. C. F. MAHADY. The effects of breathing carbon dioxide in oxygen on altitude decompression sickness. AAF School of Aviation Medicine, Proj. no. 409, Report no. 1, July 23, 1945.

(61) GREEN, H. D., R. N. LEWIS, N. D. NICKERSON AND A. L. HELLER. Blood flow, peripheral resistance and vascular tonus, with observations on the relationship between blood flow and cutaneous temperature. Am. J. Physiol. **141:** 518, 1944.

(62) GRIFFIN, D. R., S. ROBINSON, H. S. BELDING, R. C. DARLING AND E. TURRELL. The effects of cold and rate of ascent on aero-embolism. Committee on Aviation Medicine, Report no. 174, National Research Council, June 22, 1943.

(63) GRINDLAY, J. H., J. F. HERRICK AND F. C. MANN. Measurement of the blood flow of the liver. Am. J. Physiol. **132:** 489, 1941.

(64) HAGGART, G. R. AND A. M. WALKER. The physiology of pulmonary embolism as disclosed by quantitative occlusion of the pulmonary artery. Arch. Surg. **6:** 762, 1923.

(65) Haldane, J. S. and J. G. Priestley. Respiration. New Haven, Yale University Press, 1935.
(66) Haldi, J., G. Giddings and W. Wynn. Dietary control of the water content of the skin of the albino rat. Am. J. Physiol. **135**: 392, 1942.
(67) Hamman, L. Spontaneous mediastinal emphysema. Bull. Johns Hopkins Hosp. **64**: 1, 1939.
(68) Harris, M., W. E. Berg, D. M. Whitaker and V. C. Twitty. The relation of exercise to bubble formation in animals decompressed to sea level from high barometric pressures. J. Gen. Physiol. **28**: 241, 1945.
(69) Harris, M., W. E. Berg, D. M. Whitaker, V. C. Twitty and L. R. Blinks. Carbon dioxide as a facilitating agent in the initiation and growth of bubbles in animals decompressed to simulated altitudes. J. Gen. Physiol. **28**: 225, 1945.
(70) Harvey, E. N., W. D. McElroy, A. H. Whiteley, G. H. Warren and D. C. Pease. Bubble formation in animals; analysis of gas tension and hydrostatic pressure in cats. J. Cell. and Comp. Physiol. **24**: 117, 1944.
(71) Harvey, E. N. Decompression sickness and bubble formation in blood and tissues. Bull. N. Y. Acad. Med., N.S. **21**: 505, 1945.
(72) Harvey, E. N., D. K. Barnes, W. D. McElroy, A. H. Whiteley, D. C. Pease and K. W. Cooper. Bubble formation in animals; physical factors. J. Cell. and Comp. Physiol. **24**: 1, 1944.
(73) Harvey, E. M., A. H. Whiteley, W. D. McElroy, D. C. Pease and D. K. Barnes. Bubble formation in animals; Gas nuclei and their distribution in blood and tissues. J. Cell. and Comp. Physiol. **24**: 23, 1944.
(74) Hawkins, J. A. and C. W. Shilling. Nitrogen solubility in blood at increased air pressures. J. Biol. Chem. **113**: 273, 1936.
(75) Hawkins, J. A. and C. W. Shilling. Helium solubility in blood at increased pressures. J. Biol. Chem. **113**: 649, 1936.
(76) Hawkins, J. A., E. W. Shilling and R. A. Hansen. A suggested change in calculating decompression tables for diving. U. S. Nav. Med. Bull. **33**: 327, 1935.
(77) Hebb, C. O. Observations on gas bubble formation and lung damage in animals rapidly decompressed to 43,000–47,000 feet. Flying Personnel Research Committee, no. 316, June 1941.
(78) Hebb, C. O. Conditions relating to the resistance of animals to high altitudes. Flying Personnel Research Committee, no. 544, August 1943.
(79) Hebb, C. O. A critical effect of temperature in rapid decompression of rats to 44,000 feet. Flying Personnel Research Committee, no. 625, December 1944.
(80) Heller, R., W. Mager and H. v. Schrötter. Luftdruckerkrankungen. Wien: Alfred Holder, 1900.
(81) Hill, L. and M. Greenwood. The influence of increased barometric pressure in man. No. 3. The possibility of oxygen bubbles being set free in the body. Proc. Roy. Soc., London B **79**: 284, 1907.
(82) Hill, L. and M. Greenwood. Cited by L. Hill. Caisson sickness. New York, Longmans, Green & Co., 1912.
(83) Hill, L. and J. J. R. MacLeod. The influence of compressed air on the respiratory exchange. J. Physiol. **29**: 492, 1903.
(84) Hoppe-Seyler. Cited by Heller, Mager and v. Schrötter, 1857.
(85) Huggins, C. and B. H. Blockson, Jr. Changes in outlying bone marrow accompanying a local increase of temperature within physiological limits. J. Exper. Med. **64**: 253, 1936.
(86) Hutchins, H. C., A. Y. Werner, O. E. Reynolds and F. R. Philbrook. A study of aerodontalgia occurring during routine oxygen indoctrination in the low pressure chamber with a view to evolving a theory regarding its cause. U. S. M. C. Air Station, Medical Department, Quantico, Va., July 23, 1945.

(87) Ivy, A. C., A. J. Atkinson, H. Adler and W. Burkhardt. Pertaining to the effect of B²B (dextroamphetamine) and of preoxygenation plus B²B on the incidence of "bends" and "incapacitating bends and chokes" at 40,000 feet for one hour. Committee on Aviation Medicine. Report no. 113, National Research Council, December 23, 1942.

(88) Jacobs, M. H. and D. R. Stewart. Observations on the blood of albino rats following rapid decompression. Committee on Aviation Medicine, Report no. 76, National Research Council, October 1942.

(89) Jones, R. R., J. W. Crossin, E. E. Griffith, R. R. Sayers, H. H. Schrenk and F. Levy. Administration of pure oxygen to compressed air workers during decompression: prevention of the occurrence of severe compressed illness. J. Ind. Hyg. and Toxicol. **22**: 427, 1940.

(90) Jones, H. B., W. E. Myers and W. E. Berg. Gas exchange, circulation and diffusion. Committee on Aviation Medicine, Report no. 429, National Research Council, April 10, 1945.

(91) Jones, H. B., C. Tobias, W. F. Loomis, J. B. Mahoney, W. N. Sears, J. C. Larkin, J. G. Hamilton and J. H. Lawrence. An objective method for the study of the physiology of aeroemphysema and for the selection of high altitude aircrew using the radioactive isotopes of inert gases. Committee on Aviation Medicine, Report no. 81, National Research Council.

(92) Kabat, H. and C. Dennis. Decerebration in the dog by complete temporary anemia of the brain. Proc. Soc. Exper. Biol. and Med. **38**: 864, 1938.

(93) Kabat, H., C. Dennis and A. B. Baker. Recovery of function following arrest of the brain circulation. Am. J. Physiol. **132**: 737, 1941.

(94) Kabat, H. and M. Schadewald. The relative susceptibility of the synaptic terminals and of the parenchyma to arrest of the circulation of the brain. Am. J. Path. **17**: 833, 1941.

(95) Kahlstrom, S. C., C. C. Burton and D. B. Phemister. Aseptic necrosis of bone I. Infarction of bones in caisson disease resulting in encapsulated and calcified areas in diaphyses and in arthritis deformans. Surg., Gynec. and Obstet. **68**: 129, 1939.

(96) Kelman, S. Experimental emphysema. Arch. Int. Med. **24**: 332, 1919.

(97) Kenrick, F. B., C. S. Gilbert and K. L. Wismer. The super-heating of liquids. J. Phys. Chem. **28**: 1297, 1924.

(98) Kenrick, F. B., K. L. Wismer and K. S. Wyatt. Supersaturation of gases in liquids. J. Phys. Chem. **28**: 1308, 1924.

(99) Knisely, M. H., S. Gray, H. M. Peck, R. L. Nichols, L. Warner and J. A. Orcutt. The effect of elevation of a limb on the development and severity of bends pain. Committee on Aviation Medicine, Report no. 196, National Research Council, October 1, 1943.

(100) Knisely, M. H., S. Gray, H. M. Peck, R. L. Nichols, L. Warner, J. A. Orcutt and N. Anderson. Preliminary tests of the effect of intravenous aminophyllin in preventing or alleviating bends and chokes. Committee on Aviation Medicine, Report no. 195, National Research Council, October 1, 1943.

(101) Krogh, A. The anatomy and physiology of capillaries. New Haven, Yale University Press, 1929.

(102) Kuper, J. B. H. Some remarks on the formation and resolution of bubbles (with special reference to aeroembolism). Report no. 4, Studies in Aviation Medicine, Research Section of Division of Industrial Hygiene, National Institute of Health. January 3, 1942.

(103) Ladewig, P. Anoxaemic changes in the liver, with regard to the "high altitude death" of airmen. Nature **151**: 558, 1943.

(104) Lazarow, A., P. R. Patek, E. Bartosh and G. H. Scott. Observations on the capillary circulation in skeletal muscle of frogs at simulated high altitudes. Committee on Aviation Medicine, Report no. 162, National Research Council, June 15, 1943.

(105) Lund, D. W. and J. H. Lawrence. An hypothesis as to a cause of "bends" pain with observations on massage at high altitude. Committee on Aviation Medicine, Report no. 404, National Research Council, January 10, 1945.
(106) Lund, D. W., J. H. Lawrence and L. B. Lawrence. Latent neurological manifestations following decompression. Occupational Medicine 1: 75, 1946.
(107) Macklin, C. C. Pneumothorax with massive collapse from experimental local over-inflation of the lung substance. Can. M. A. J. 36: 414, 1937.
(108) Macklin, C. C. Transport of air along sheaths of pulmonic blood vessels from alveoli to mediastinum. Arch. Int. Med. 64: 913, 1939.
(109) McElroy, W. D., A. H. Whiteley, K. W. Cooper, D. C. Pease, G. H. Warren and E. N. Harvey. Bubble formation in animals; physiological factors: the rôle of circulation and respiration. J. Cell. and Comp. Physiol. 24: 273, 1944.
(110) McElroy, W. D., A. H. Whiteley, G. H. Warren and E. N. Harvey. Bubble formation in animals; relative importance of carbon dioxide concentration and mechanical tension during muscle contraction. J. Cell. and Comp. Physiol. 24: 133, 1944.
(111) McGuire, J. and W. B. Bean. Spontaneous interstitial emphysema of the lungs. Am. J. Med. Sci. 197: 502, 1939.
(112) Mediastinal emphysema (pneumomediastinum) following explosive decompression of humans; report of two cases. Jan. 1, 1945, Memorandum Report of Aero Medical Laboratory, Wright Field, Dayton, Ohio, no. TSEAL-3-695-29-I.
(113) Meringi, G., G. H. v. Schottz and J. Gyurik. Über die Wirkung von Ammonium Chloride auf die Hohenfestigkeit des Menschen. Luftfahrtmed. 5: 102, 1941.
(114) Mitchell, D. F. Aerodontalgia. Bull. U. S. Army Med. Dept., no. 73, 62, 1944.
(115) Morales, M. F., E. N. Rathbun, R. E. Smith and N. Pace. Studies on body composition: theoretical considerations regarding major body tissue components, with suggestions for application to man. J. Biol. Chem. 158: 677, 1945.
(116) Morales, M. F. and R. E. Smith. On the possible determination of gross human body composition by the use of radioactive inert gases. Research Project X-43, Report no. 4, Naval Medical Research Institute, August 1, 1945.
(117) Morales, M. F. and R. E. Smith. The physiological factors which govern inert gas exchange. Bull. Math. Biophys. 7: 99, 1945.
(118) Morales, M. F. and R. E. Smith. On the theory of blood-tissue exchanges: III. Circulation and inert gas exchanges at the lung with special reference to saturation. Bull. Math. Biophys. 6: 141, 1944.
(119) Morales, M. F. and R. E. Smith. A note on the physiological arrangement of tissues. Bull. Math. Biophys. 7: 47, 1945.
(120) Morey, J. B. and M. C. Sosman. Spontaneous mediastinal emphysema, with report of case associated with spontaneous pneumothorax. Radiology 32: 19, 1939.
(121) Osipov, M. O. and V. F. Lashkov. A contribution to the pain problem at high altitude. Transl. by S. Kohn, R. C. A. M., July 15, 1940.
(122) Oudard. Accidents de décompression: Relation d'autopsie. Arch. de Méd. et Pharm. Nav. 95: 63, 1911.
(123) Pace, N. Equations for the estimation of total body fat and total body water from the solubility of inert gases in the body. Research Project X-191, Report no. 4, Naval Medical Research Institute, September 25, 1945.
(124) Pappenheimer, J. R. and J. P. Maes. The effects of vasoconstriction on the apparent viscosity of blood flowing through the hindlimb of the dog. Federation Proc. 1: 65, 1942.
(125) Patek, P. R., A. Lazarow, E. Bartosh and G. H. Scott. Observation on living periosteum and peritoneum at simulated high altitudes. Committee on Aviation Medicine. Report no. 244, National Research Council, January 17, 1944.
(126) Pease, D. C., L. R. Blinks and E. A. Reed. Bubble formation in decompressed

animals. VII. Physical factors in bubble formation. Committee on Aviation Medicine. Final Report OEMcmr-193, National Research Council, October 15, 1944.

(127) Phemister, D. B. Changes in bones and joints resulting from interruption of circulation. II. Non-traumatic lesions in adults with bone infarction; arthritis deformans. Arch. Surg. **41:** 1435, 1940.

(128) Piccard, J. Aeroemphysema and the birth of gas bubbles. Proc. Staff. Meetings, Mayo Clinic **16:** 700, 1941.

(129) Pigott, J. P. Dental pain at high altitudes; origin and treatment. Interim Report OEMcmr-38, National Research Council, October 23, 1944.

(130) Polak, I. B. and B. H. Adams. Traumatic air embolism in submarine escape training. U. S. Nav. Med. Bull. **30:** 165, 1932.

(131) Polak, I. B. and C. L. Tibals. A fatal case of caisson disease following a dive of short duration to a depth of thirty feet. U. S. Nav. Med. Bull. **28:** 862, 1930.

(132) Pudenz, R. H. Personal communication.

(133) R.A.F. Institute of Aviation Medicine, Physiological Research Unit. Farnsborough, England. Observations on Decompression Sickness in Man. Flying Personnel Research Committee Rept. no. 267. March, 1941.

(134) Rathbun, E. N. and N. Pace. Studies on body composition. I. The determination of total body fat by means of the body specific gravity. Research Project X-191, Report no. 1, Naval Medical Research Institute, August 7, 1944.

(135) Ratnoff, O. D. The absence of roentgenographically demonstrable bony changes at the hip joint in subjects exposed to simulated high altitudes. School of Aviation Medicine, Randolph Field, Texas, Report no. 201, November 12, 1943.

(136) Recommendation for the handling of reactions following altitude chamber flights. A. A. F. School of Aviation Medicine, Randolph Field, Texas, Project no. 217, Report no. 1, December 30, 1943.

(137) Reed, E. A. and L. R. Blinks. Bubble formation in decompressed animals. V. The relation of temperature and exercise to bubble formation in rats, and in tourniqueted legs of rabbits and goats. Final Report no. 377, National Research Council, October 15, 1944.

(138) Reed, E. and L. R. Blinks. Bubble formation in decompressed animals. VI. Vasoconstriction and the relation of the vascular bed to bubble formation in frogs. Committee on Aviation Medicine, Report no. 379, National Research Council, October 15, 1944.

(139) Rendich, R. A. and L. A. Harrington. Roentgen findings in caisson disease of bone, with case reports. Radiology **35:** 439, 1940.

(140) Richardson, H. F., B. C. Coles and G. E. Hall. Experimental gas embolism. I. Intravenous air embolism. Canad. M. A. J. **36:** 584, 1937.

(141) Ryder, H. W., G. L. Engel, J. Romano, J. P. Webb, M. A. Blankenhorn, E. B. Ferris and W. E. Brown. An assay of dextro-amphetamine for its protective value in decompression sickness. Committee on Aviation Medicine, Report no. 112, National Research Council, January 28, 1943.

(142) Schmidt, C. F., S. S. Kety and H. H. Pennes. The gaseous metabolism of the brain of the monkey. Report no. 389, OEMcmr-28, National Research Council, September 1, 1944.

(143) Schultz. Cited by Heller, Mager and v. Schrötter, 1878.

(144) Sendroy, J., Jr., R. T. Dillon and D. D. Van Slyke. Studies of gas and electrolyte equilibrium in the blood. XIX. The solubility and physical state of uncombined oxygen in the blood. J. Biol. Chem. **105:** 597, 1934.

(145) Sharpless. Cited by Heller, Mager and v. Schrötter.

(146) Shaw, L. A., A. R. Behnke, A. C. Messer, R. M. Thomson and E. P. Motley. The equilibrium time of the gaseous nitrogen in the dog's body following changes of nitrogen tension in the lungs. Am. J. Physiol. **112:** 545, 1935.

(147) SJÖSTRAND, T. On the principles for the distribution of the blood in the peripheral vascular system. Skand. Arch. f. Physiol. 71: suppl., 1935.
(148) SMITH, H. The physiology of the kidney. New York: Oxford University Press, 1937.
(149) SMITH, J. J. Effects of explosive decompression on animals. War Department, Air Corps Materiel Division, Report no. EXP-M-54-653-34D, May 20, 1942.
(150) SMITH, R. E. AND M. F. MORALES. On the theory of blood tissue exchange: I. Fundamental equations. Bull. Math. Biophys. 6: 125, 1944.
(151) SMITH, R. E. AND M. F. MORALES. On the theory of blood tissue exchanges: II. Applications. Bull. Math. Biophys. 6: 133, 1944.
(152) SNYDER, C. D. Recent advances in knowledge of the liver. Physiol. Rev. 22: 54, 1942.
(153) STOEL, G. Zellforsch. 3: 91, 1925, cited from Krogh.
(154) SWINDLE, P. F. ET AL. The possible relationship between intravascular agglutination of erythrocytes and decompression sickness. Committee on Aviation Medicine, Report no. 178, National Research Council, August 13, 1943.
(155) THOMAS, S. AND O. L. WILLIAMS. High altitude joint pains: their radiographic aspects. Committee on Aviation Medicine, Report no. 395, National Research Council, December 11, 1944.
(156) THUDICHUM, T. L. W. Die Chemische Konstitution des Gehirns der Menschen und der Thiere. Tübingen: F. Putzcker, 1901.
(157) TOBIAS, W. F., W. F. LOOMIS, F. C. HENRY, W. R. LYONS, H. B. JONES, W. N. SEARS, S. F. COOK, J. B. MOHNEY, J. G. HAMILTON AND J. H. LAWRENCE. Circulation and decompression sickness. Committee on Aviation Medicine, Report no. 144, National Research Council, June 7, 1943.
(158) TROWELL, O. A. Liver vacuoles and anoxia. Nature 151: 730, 1943.
(159) TROWELL, O. A. Histological changes in the lungs of rabbits decompressed to 40,000–47,000 feet. Flying Personnel Research Committee, no. 317, June 1941.
(160) TROWELL, O. A. A histological examination of the spinal cord of animals rapidly decompressed to 40,000–45,000 feet. Flying Personnel Research Committee, no. 345, August, 1941.
(161) TURPIN, F. H., W. F. LOOMIS, J. H. LAWRENCE, H. B. JONES AND C. A. TOBIAS. Solubilities of gases in water and oils. Report no. 455, National Research Council, July, 1945.
(162) TUREEN, L. L. AND J. B. DEVINE. The pathology of air embolism. J. Missouri State Med. Assn., 33, 141, 1936.
(163) TWYNAM, G. E. A case of caisson disease. Brit. Med. J. 1: 190, 1888.
(164) UNDERWOOD, N. AND J. T. DIAZ. A study of the gaseous exchange between the circulatory system and the lungs. Am. J. Physiol. 133: 88, 1941.
(165) VAN ALLEN, C. M., G. L. NOCOLL AND W. M. TUTTLE. Lung changes after occlusion of pulmonary artery branches by embolus and by ligature. Yale J. Biol. and Med. 2: 363, 1929–30.
(166) VAN RENSSELAER, H. The pathology of the caisson disease. Med. Rec. 40: 141, 178, 1891.
(167) VAN SLYKE, D. D., R. T. DILLON AND R. MARGARIA. Studies of gas and electrolyte equilibria in blood. XVIII. Solubility and physical state of atmospheric nitrogen in blood cells and plasma. J. Biol. Chem. 105: 571, 1934.
(168) VERNON, H. M. The solubility of air in fats and its relation to caisson disease. Proc. Roy. Soc. London B 79: 366, 1907.
(169) VIGUIER AND G. JEAN. Gaseous embolism of the gluteal artery (produced by decompression). Bull. Acad. de Med. 80: 377, 1918.
(170) VILLARET, M., R. CACHERA AND R. FAUVERT. L'embolie gazeuse cérébrale; ses effets circulatoires locaux. C. R. Soc. de Biol. 125: 108, 1937.

(171) v. LEYDEN. Cited by Heller, Mager and v. Schrötter, 1877.
(172) WAGNER, C. E. Observations of gas bubbles in pial vessels of cats following rapid decompression from high pressure atmospheres. J. Neurophysiol. 8: 29, 1945.
(173) WARREN, C. O., JR. The oxygen consumption of rabbit bone marrow in relation to its morphology. Am. J. Physiol. 110: 61, 1934.
(174) WEBB, J. P., E. G. FERRIS, JR., L. ENGEL, J. ROMANO, H. W. RYDER, C. D. STEVENS AND M. A. BLANKENHORN. Radiographic studies of the knee during bends. Committee on Aviation Medicine, Report no. 305, National Research Council, May 8, 1944.
(175) WEARN, J. T. Morphological and functional alterations of the coronary circulation. The Harvey Lectures, 243, 1939–40.
(176) WEINBERGER, L. M., M. H. GIBBON AND J. H. GIBBON, JR. Temporary arrest of the circulation to the central nervous system. Arch. Neurol. and Psychiat. 43: 615, 1940.
(177) WHITAKER, D. M., L. R. BLINKS, W. E. BERG, V. C. TWITTY AND M. HARRIS. Muscular activity and bubble formation in animals decompressed to simulated altitudes. J. Gen. Physiol. 28: 213, 1945.
(178) WHITAKER, S. R. F. AND F. R. WINTON. The apparent viscosity of the blood flowing in the isolated hind limb of the dog, and its variation with corpuscular concentration. J. Physiol. 78: 339, 1933.
(179) WHITELEY, A. H., W. D. MCELROY, G. H. WARREN AND E. N. HARVEY. Bubble formation in animals; denitrogenation. J. Cell. and Comp. Physiol. 24: 257, 1944.
(180) WINDLE, W. F., R. F. BECKER AND A. WEIL. Alterations in brain structure after asphyxiation at birth. An experimental study in the guinea pig. J. Neuropath. and Exper. Neurol. 3: 224, 1944.
(181) WOLFF, H. G. The cerebral blood vessels—anatomical principles. Assn. Res. Nerv. and Ment. Dis. 18: 29, 1938.
(182) YASKIN, J. C. AND M. W. THORNER. The effects upon the cerebral cortex of altitude chamber anoxia. Trans. Am. Neurol. Assn. 69: 88, 1943.
(183) ZETTERSTRÖM, A. Deep sea diving with synthetic gas mixtures. Report from the Swedish Admiralty, 1944.
(184) ZOGRAFIDI, S. Contribution à l'étude des accidents de decompression chez les plongeurs à scaphandre. Rev. de Med. 27: 159, 1907.
(185) ZUNTZ, N. Zur Pathogenese und Therapie der durch rasche Luftdruckänderung erzeugten Krankheiten. Fortschritte der Med. XV Jahr., 632, 1897.

THE GENESIS OF THE ELECTROCARDIOGRAM

LOUIS N. KATZ

The Cardiovascular Department, Research Institute, Michael Reese Hospital, Chicago, Ill[1]

A review of the significance of the electrocardiogram may present an extensive noncommittal report of a voluminous literature or a summary of the author's viewpoint based on a fair evaluation of the pertinent studies of others and on his own experiences and deductions. The latter approach was chosen. Several excellent reviews have appeared (10, 13, 27, 61, 153, 155, 162, 163, 294, 299, 309, 314, 356, 367) since the last PHYSIOLOGICAL REVIEW article written by the author in 1928 (161). The present review will concern itself only with the theoretical aspects of the electrocardiogram. The empirical interpretations of the electrocardiogram which have been found so useful in clinical practice will not be discussed since they have been dealt with in extenso elsewhere (164).

Any careful survey of the articles appearing on the subject must impress the unbiased reader with the fact that the manifold theories on the genesis of the electrocardiogram are in apparent contradiction with one another. He may perforce in desperation apply electrocardiography without regard to theory, leaving the answer to the latter to the future. It is very fortunate that the clinical usefulness of the electrocardiograph has suffered so little from lack of agreement on its theoretical aspects. While some workers continue to proclaim the need for an understanding of the theory in order to make electrocardiography useful for clinical practice, the electrocardiograph has actually been very intensively and profitably used without any but the most elementary theoretical background. It may be stated in all fairness that the attempt to instill a theoretical grounding into the ordinary clinician who wishes to use this tool has served too often to confuse rather than to assist him. This is so because the background of physics, mathematics and physiology which the ordinary electrocardiographer possesses is too superficial for a critical evaluation of the facts and fancies presented. Too many presentations of the subject make untenable assumptions of the constancy of certain variables in order to arrive at a "simple" mathematical synthesis of the subject. Some recognize the assumptions made at the start but too often lose sight of them in later developments, especially when the subject is taken over by enthusiastic and worshipful disciples. Having once established a point of view, investigators often continue to expand it until the superstructure of the theory becomes too heavy for the narrow pivot of fact upon which it rests. This attempt to make the interpretation of data yield more precise quantitative information than is justified by the crude initial assumptions, is the basic fault in many instances. It is hoped that this will become apparent as this review is developed.

What has just been stated does not apply to the electrocardiographic interpretation of cardiac arrhythmias. Here knowledge of the normal and abnormal

[1] The department is supported in part by the Michael Reese Research Foundation.

physiology of impulse origin and impulse spread, of the refractory period of the heart and of its syncytial character is of vital importance. The more the electrocardiographer understands these fundamental properties, the more rational become his interpretations. However, this review will not deal with the arrhythmias except in so far as such knowledge helps account for the creation of electric currents by the heart.

Factors Involved in the Genesis of the Electrocardiogram. The genesis of the electrocardiogram is but a special and perhaps more complex case of cellular electrophysiology. Its features are therefore basic in general physiology. However, any analysis of the electrocardiogram must concern itself with the fact that the electrodes used for recording currents are applied to the body surface at a distance from the heart. Even when one of the electrodes is placed over the precordium it is still a distant electrode. Consequently, four principal factors are involved in this analysis:

a. The genesis and typical time course of the electric states of the heart during activity and injury.

b. The laws governing the spread of current within a volume conductor.

c. The influence exerted by the characteristics of the body as a conductor of electricity.

d. The modifications introduced by variations of the anatomy and physiology of the heart.

The Genesis and Typical Time Course of the Electric States of the Heart during Activity and Injury. A. The electric state of resting heart muscle. It is now well established that (a) the composition of living cells differs from that of the ambient extra-cellular fluid (72), and (b) that the surface membrane of cells is semi-permeable. It follows, therefore, that the resting cell surface is electrically polarized; this polarity is such that the charge on the inside is negative and that on the outside is positive. It is not unlikely that other electric charges exist at interfaces within cells as complex as that of the heart cell syncytium, but there is no reason to believe that these are involved in the genesis of the electric current recorded in the electrocardiogram. The electric characteristics of the cell surface may not be as simple as suggested above. For example, Cole (59, 59A) has shown that during flow of current across a surface membrane it possesses electric properties of impedance, capacitance and inductance. The classical hypothesis of the polarized membrane, however, is adequate for an understanding of the genesis of the electrocardiogram. The magnitude of the potential across the resting cell membrane was found to be of the order of 100 millivolts in certain cells in which it can be measured (62, 63). However, at rest since, all parts of the cell membrane are maintained uniformly polarized, no potential differences will exist within the cell nor in the medium surrounding the cell, and no electric currents will flow.

B. The alterations of the electric state during activity. It is now generally agreed that during activity of a living cell, the permeability of its surface is altered, the impedance across the surface changes, the ions which are responsible for the polarized state of the resting cell migrate and that the cell becomes

depolarized. Recently Cole (63) observed that following depolarization a temporary reversal of polarization occurs. If confirmed, this would show that the process of local activation is not as simple as has been hitherto conceived. This can be neglected for our purposes since the magnitude of potential differences and current involved is much smaller than that due to depolarization of the "resting" charges.

The process of depolarization in any one region is not instantaneous but of measurable duration. This time plus that occupied in the propagation of the depolarization process throughout the cell surface determine the duration of the process of depolarization. Since during this period the electric state of the cell membrane is not uniform, differences in potential are established within the cell and also in any conducting medium surrounding the cell, and electric currents will flow while these potential differences exist.

When the depolarization is completed the cell may remain in the depolarized state for a variable time. During this time there is no reason to believe that any potential differences exist and hence no electric currents will flow. Thus the cell is electrically quiescent even though it may, as in the case of a contractile organ like the heart, be undergoing changes in tension and in length.

The process of repolarization, which follows soon after depolarization, is definitely slower than the process of depolarization. Further, the pattern of depolarization need not have the same topographical distribution as that of the depolarization preceding it. The evidence is clear that in the case of the heart the duration of electrical systole, that is, the period from the beginning of depolarization until repolarization is completed, is unequal in various sites of the ventricles (cf. 164). Since the electric state of the cell is not uniform during repolarization, potential differences are again established and electric current flows once more within the cell and in any conducting medium surrounding the cell. However, the potential differences are now not as great as during depolarization since the process of repolarization is slower. If the topography of repolarization in the cell were identical with that of depolarization and if the duration of electrical systole were precisely the same in all parts of the cell (unlikely occurrences in the normal ventricles), then the potential differences during repolarization would be opposite in sign to those in the stage of depolarization and the area subtended by it would equal that subtended during depolarization.

The theory outlined above is widely accepted by students of electrocardiography. It is also agreed that these electric phenomena are not manifestations of the mechanical responses of heart muscle, since the time pattern of their occurrence is not identical with that of the mechanical phases (175, 339, 341). The details of the energetics and chemical factors involved in the process leading to electric manifestations (223, 224) are not yet fully understood, but are not the concern of a review which deals only with the electric phenomena themselves. For the genesis of the electrocardiogram it suffices to recognize the following four states of the cell, viz.: *a*, the polarized state of rest; *b*, the state of depolarization—during activation; *c*, the depolarized state of the active cell,

and *d*, the stage of repolarization—during restitution (or recovery). Further, it is important to appreciate that only during the process of depolarization and repolarization do potential differences appear and electric currents flow. The electrogram thus records a double response during the heart cycle when the electrodes are placed on two uninjured parts of the heart's surface, an initial large and brief deflection and a longer and smaller final deflection, with an interval between them. This is the genesis of the P and T_A waves of the auricles and of the QRS complex, S-T segment and T wave of the ventricles in the electrocardiogram.

C. The alterations of the electric state during injury. Injury involving a part of the cell may do a number of things (cf. 318):

1. Injury may completely depolarize a part of the surface of the affected cell, and hence this region does not share in the activity of the rest of the cell. There will be potential differences during the resting polarized state of the uninjured part of the cell, because the cell membrane will not be uniformly polarized; an injury current of rest will therefore be present. During depolarization of the uninjured part this potential difference will disappear since the cell becomes completely depolarized, only to reappear again during repolarization of the uninjured part. Thus there will be an injury current of rest, but no injury current of activity. An injury of this type will result in a monophasic electrogram during the heart cycle if electrodes leading from an uninjured and an injured part of the heart's surface are connected to a recording galvanometer. The S-T deviations in the electrocardiogram caused by injury can be accounted for on this basis. The application of potassium to the surface of the heart appears to operate in this way (23).

2. Injury may cause a partial, rather than a complete, depolarization of the region of the cell affected, but still leave it responsive. During the resting stage, potential differences between injured and uninjured regions will exist but will be of smaller magnitude than in the case of complete depolarization of the injured part. During activation the injured part will also respond and so the potential differences in the cell will disappear. There will thus be a monophasic curve of smaller magnitude than in the preceding situation. In this case also there will be only an injury current of rest and no injury current of activity. This type of injury is thus a variation of the preceding one.

3. Injury while causing only a partial depolarization of the part of the cell affected may also make this area completely or partially unresponsive. During the resting stage the situation would be similar to that of the preceding cases in producing an injury current of rest, but the injured part of the cell would retain its partially polarized state during activity. Thus, when the rest of the cell is depolarized, potential differences of opposite directions to those previously existing would develop. There would thus be not only an injury current of rest but also one of activity. Again a monophasic curve would be recorded when an injured and an uninjured part of the cell are connected to a galvanometer, and S-T deviations would result in the electrocardiogram. But the monophasic curve could cross the null point, once during depolarization and again during

repolarization. This appears to be the common effect of injury produced with the suction (89, 93) or the pressure electrode (318).

4. Injury may result in unresponsiveness of the injured region without affecting its resting polarization, or injury may lead to a region of block preventing the impulse from reaching some part of the cell. Under these circumstances no injury current would flow during the resting state of the cell. But during the activity stage, the injured, or isolated, part of the cell would retain its surface polarization while the rest of the cell was depolarized. There would thus be an injury current of activity but no injury current of rest. Again a monophasic curve would be recorded when an injured and uninjured (or blocked off) part of the cell were connected to a galvanometer, but it would not cross the null point. It would cause S-T deviations in the electrocardiogram. Cold appears to operate in this manner (23).

5. Finally injury may cause a delay and protraction of the process of repolarization in a responsive and fully polarized or partially depolarized part of the cell. This would lead to an injury current of repolarization which would modify the appearance of the final deflection of the electrogram obtained by connecting the injured and uninjured part of the cell (318). This would account for the T wave of injury in the electrocardiogram. If the injured region were completely polarized at rest, there would be no injury current of rest or activity, but only an injury current of repolarization. If the injured part of the cell were partially depolarized there would be not only an injury current of repolarization but also an injury current of activity. It would account for the frequent combination in the electrocardiogram of S-T deviations with oppositely directed T waves. A variant of this would be the case in which the injured area was partially depolarized, partially unresponsive and also had a protracted period of repolarization. Here there would be an injury current at rest, an injury current during activity and an injury current of repolarization. Injury currents of repolarization may occur in areas of the cell different from those giving rise to the other types of injury currents; usually the two areas are adjacent to each other (318). The amount of delay and protraction in the repolarization of the injured area may vary, accounting for different degrees of injury currents of repolarization and various magnitudes and durations of injury T waves.

The exposition above will account for the genesis of the S-T and T wave deviations which arise as a result of injury. It remains to point out that when a part of the injured cell is partially or wholly unresponsive this will alter the appearance of the initial deflection of the electrogram obtained in connecting an injured with an uninjured area and also of the QRS deflection of the electrocardiogram since the orientation of the potential differences during the process of depolarization of the cell are altered. (Some of the T wave alterations in injury may be on a similar basis; namely, reorientation of the potential differences during repolarization because of absence of response of the injured area.)

The living cell ordinarily does not permit the state of injury to persist but tends to restore the normal state. Sometimes, however, the processes giving rise to injury balance the restitution processes, and the injury state is maintained

for long periods of time. The newly formed part of the polarized surface may enclose the injured area which then is restored to the cell, or it may develop at the junction between the injured and uninjured parts of the cell, the former thus becoming dead tissue. This development of a newly formed surface of polarization causes the injury pattern to change with the passage of time. It also accounts for the tendency for the effects of injury to change from injury currents during rest and/or activity, causing S-T deviations, to injury currents during repolarization, causing T changes. On the other hand, if a mild injury became more severe, the reverse would take place. As soon as recovery from injury is complete, the injury S-T deviations and T wave alterations disappear. Thus it may be stated categorically that S-T deviations and T wave alterations of injury indicate the persistence of injury currents, whereas QRS alterations due to injury need not disappear since they often remain when the injured area has been replaced by necrotic or scar tissue.

Considerable controversy has arisen as to whether it is the injured or the uninjured part of the cell which is responsible for the injury currents (56, 61, 80, 81, 89, 94, 102, 103, 126, 156, 168, 215, 225, 307, 308, 309, 310, 318, 367, 369). The discussion given above shows that both areas are involved. If no injured area is present, no injury current will develop; and likewise if *no* part of the cell remains uninjured, no injury current will occur. Injury currents signify that part of the cell is injured, but not all of it. The source of the current is obviously ultimately referable to that part of the cell which is polarized. Depending on circumstances this may be the uninjured or the injured part, or both, and this can be worked out for the various effects of the injury, as outlined above. Actually, experience shows that the onset, duration and end of the monophasic curve depends on the electrode placed over the uninjured area rather than on that placed over the injured area (156, 309; also cf. 80). The modification introduced when indirect leads, or leads involving one electrode on the surface of the heart and the other at a distance are employed, will be discussed below.

The attempt to read into the monophasic curve a series of phenomena as has recently been done (289, 290, 291, 292) has rightly been criticized as a misinterpretation of artefacts in the record (11A, 80). Even with the best technique the curves obtained may not be pure injury currents, but may contain some distortions, since the electrode over the injured area may pick up currents from uninjured areas in the neighborhood of the injury.

D. The syncytial character of the heart. The heart is composed of a complex of spiral muscles arranged in sheaths and in bundles (262, 279). The muscle bundles of the auricles are distinct from those of the ventricles, and separated from them by a fibrous tissue sheet at the auriculo-ventricular junction to which the heart valves are attached. In the mammal and bird there is ordinarily only one muscular bridge between the two, consisting of neomorphic specialized muscular tissue. This so-called auriculo-ventricular bridge is made up of the auriculo-ventricular node and the common auriculo-ventricular bundle. While some recent workers have questioned the specialized character of this bridge in the dog, monkey and man, because the histological differentiation from ordinary

heart muscle is not as striking as in other species (105, 106, 107, 108), the overwhelming mass of evidence shows its anatomical existence (37, 38, 58, 60, 64, 65, 66, 67, 68, 178, 198, 201, 208, 209, 210, 235, 270, 272, 284, 293, 294, 327, 332, 333, 334, 338, 378, 379) and, more important, shows it to be the sole bridge which permits the transfer of the impulse from the auricles to the ventricles. The recent attempt to revive the hypothesis of the neurogenic transfer of the impulse from the auricles to the ventricles (105, 106, 107, 108) has not presented new evidence. The controversy raging a quarter of a century ago was finally settled in favor of the myogenic hypothesis of impulse transmission, and the evidence today still favors it (198, 294, 327). Of course, it is true that in mammals there may occasionally exist other muscular connections between the auricles and ventricles, like the bundle in the lateral wall of the right heart found by Kent in some herbivora. In fact, a type of electrocardiogram with an abnormally short P-R interval and an abnormally long QRS duration (the Wolff-Parkinson-White syndrome) has recently been explained on the basis of such a congenital by-pass. This by-pass, in various locations, has actually been demonstrated at necropsy in such cases (240, 377). Further, it must be admitted that the auriculo-ventricular connective tissue partition is indented by ventricular and auricular muscle so that it has corrugated surfaces. To elucidate the rôle played in transmission by muscle bundles penetrating the auriculo-ventricular fibrous partition, a systematic survey of human hearts with serial wax model preparations is needed. It appears unlikely that such a survey will alter the classical, widely accepted views of the nature and function of the auriculo-ventricular muscular bridge.

The auricular and ventricular musculature, therefore, each form a single unit, linked together by a narrow muscular bridge. Each may be considered as a syncytium. While it may be argued on histological grounds that some of the cross-striations divide the muscle into distinct cellular compartments, from the point of view of the genesis of the electrocardiogram the muscle fibers of the auricles and those of the ventricles may be considered as each constituting a single though extremely complex syncytial cell. On this hypothesis there is fairly complete agreement.

This concept of the syncytial nature of heart muscle is important in explaining the continuity of the impulse transmission over the various chambers of the heart. It readily explains why an injured region within the heart muscle which does not border on the endocardium or epicardium creates no external electric field. It explains why an impulse spreading radially in all directions from a point within the heart muscle does not create an external electric field until it reaches the endocardial, epicardial or the auriculo-ventricular fibrous partition surfaces. If the myocardium were composed of discrete electrically polarized cells rather than a single syncytial cell each for the ventricles and auricles, then injury entirely within the myocardium, unless it were such as to affect the entire surface of the cells involved, should on some occasion cause an external electric field. On the basis of discretely discharging multiple cells, it would be difficult to account for the deflections occurring during the cycle of cardiac activity. The potential differences occurring during depolarization and repolarization exist only as long

as the cell throughout its surface is not in an identical electric state. There would be no reason to find a set of deflections for auricular activity distinct from those of ventricular activity since it is difficult to see how the spread of depolarization and repolarization over large numbers of discrete cells could summate in such a way as to give the large deflections of the electrical record obtained from the heart. Such problems in explaining the electrocardiogram are avoided by accepting the view that the heart is made up of two cells, one composing the auricles, the other the ventricles.

E. Other sources of electric currents in the heart. From time to time other sources of electricity have been invoked to account for the various deflections of the electrocardiogram (196, 202, 217; cf. 164). Frictional electricity, for example, has been suggested to explain the T wave, although the evidence has never appeared convincing. While movements of the heart and changes in the electrical impedance (253, 288) during the heart cycle may modify the distribution of the electric currents, they cannot be considered as a cause for the electric forces which are created during the heart cycle. The P and T_A waves of the auricles, and the QRS deflection, the S-T segment and the T wave of the ventricles are adequately accounted for on the basis of the polarization hypothesis developed above. One dilemma in the account of the deflections of the electrocardiogram is the causation of the U wave. While a number of suggestions have been made to account for this deflection (138, 226, 227, 243, 312), its exact cause remains to be determined. It is not unlikely that it is similar in character to the after-discharges found in nerve fibers, which occur also in the heart (45, 311).

F. Topography of impulse origin and spread. The various deflections of the electrocardiogram depict the time course of the processes of depolarization and repolarization in the auricles and ventricles. The meaning of this time pattern will become clearer once the spatial pattern of impulse origin and spread as determined by their topography in the heart is clearly understood. In this review only a bare outline of this topography of the heart will be attempted, based on the classical view which remains ascendant despite numerous attacks over the past four decades. The mechanisms of impulse initiation and those of impulse propagation can not be considered here, nor can any detailed account be given of the anatomy of the muscular elements involved.

The impulse originates in the sinus node located in the sulcus terminalis of the right auricle, or more precisely in the head of this node. The sinus node is a neomorphic development of specialized musculature in which the property of rhythmic discharge is highly developed. It is also a tissue of slow conductivity. With the discharge of the pacemaker in the node, the impulse spreads slowly over this node and then enters the auricular musculature by way of transitional muscle fibers. In the auricles it appears to travel faster and following the intricate branching of the auricular musculature ultimately stimulates, that is depolarizes, all parts of the auricles. Soon after depolarization, the depolarized auricular musculature repolarizes with a topographical spread more or less similar to that followed by depolarization so that the ordinary electrical record of

auricular activity shows two deflections, the P an T_A waves, opposite to each other.

In due course, the impulse in the auricles arrives at the auriculo-ventricular bridge described above and, passing through transitional fibers, it enters the auriculo-ventricular node. This node is also a neomorphic development of specialized muscle with anatomical and physiological properties resembling those of the sinus node. The claim that a special musculature bridge connects the two nodes (241, 242) appears to be unfounded except as a rare anomaly. Because the impulse travels slowly in nodal tissue, a delay in the transmission of the impulse to the ventricles is found. Once the impulse has passed the auriculo-ventricular node, it enters directly into the common auriculo-ventricular bundle, then its bundle branches and the terminal net-work of specialized muscular fibers. This latter system also is a neomorphic development of specialized muscle. In the common auriculo-ventricular bundle the character of the tissue quickly changes to the characteristic Purkinje fibers which are capable of rapid transmission of the impulse. It has been shown that except for the upper part of the interventricular septum, the terminal net-work penetrates all parts of the ventricular musculature (1, 2, 4). The bundle branches and the Purkinje net-work lie under the endocardium but the net-work also penetrates to the epicardium in all parts of the ventricles. The impulse therefore spreads simultaneously over many parts of the ventricles by means of this rapidly conducting terminal net-work, but in each locality arriving at the endocardium before reaching the epicardium. An objection to this view raised by Robb (263, 267, 271) has recently been withdrawn by her (264). In each part of the ventricles the impulse passes from the terminal Purkinje net-work via transitional fibers to the ventricular musculature proper. Thus the entire ventricular musclature is depolarized by a multitudinous series of quickly succeeding local impulses rapidly disseminated by the terminal net-work. This rapid spread makes the time of depolarization of the more massive ventricles of the same order of duration as that of the auricles, i.e., the QRS duration is of the same order as the P duration. Repolarization of the ventricles does not follow their depolarization as quickly as is the case in the auricles, so that a depolarized state expressed by the S-T segment is recorded in the ventricles. The electrical systole of the ventricles is much longer than that of the auricles. Repolarization of the ventricles does not follow the same topographic pattern as depolarization since the duration of electrical systole in the various parts of the ventricles shows considerable variation. Hence in many of the leads of the electrocardiogram the T wave is not opposite to the direction of the QRS deflection.

Many attempts have been made to depict the arrival of the impulse over the surface of the ventricles with various types of electrical connections in a number of species (2, 24, 88, 90, 91, 104, 123, 127, 128, 198, 199, 237, 257, 258, 271, 336, 343, 346). The data obtained with electrodes on various injured parts of the heart connected to a distant electrode can be dismissed as not suitable. Observations with the differential electrode, with electrodes on various uninjured parts of the heart connected to a distant electrode, and the more extensive data of

Harris (127, 128) with adjacent punctuate electrodes on various spots of the heart are more pertinent. The discrepancies in the observations of the various investigators can be explained on the basis of differences in the state of the exposed hearts studied and the adequacy of the criteria used to determine the arrival of the impulse locally. The results are in agreement, however, in showing 1, that stimulation of the external surface of the heart is not simultaneous and that the impulse often reaches regions far apart more nearly simultaneously than adjacent parts, and 2, that the total time occupied in stimulating the surface of the ventricles is much shorter than the duration of the QRS complex. The latter is partly due to the time occupied in stimulating the regions of the ventricles not bordering the external surface and partly due to the fact that the depolarization process in any region occupies time, which is significant when compared with the transmission time over the ventricles.

It is not feasible as yet, as some have attempted, to depict the exact details of the spread of the impulse through the auricles or ventricles. We are in ignorance of too many factors to be able to do this. However, some general comments on the direction of impulse spread in the average normal human heart can be made. In general in the auricles, the impulse spreads ventrally, caudally and to the left. This is adequately explained on the basis of the location of the sinus node in the upper right portion of the auricles and the assumption that there is a fairly uniform rate of spread through the auricles. There is no reason to assume the presence of specialized conducting paths or of synaptic-like junctions (256, 257, 258), nor is there need to assume that the impulse preferentially follows the spiral sheets and bundles of muscle.

In the case of the ventricles the situation is more complex. The impulse invades the septum before the outer walls of the ventricles. In the septum the muscle fibers proper appear to be reached first in two sites, one beneath the endocardium of the right and the other beneath that of the left ventricle, both sites being some distance below the auriculo-ventricular fibrous junction. The impulses from these two sites soon meet somewhere between the two endocardial surfaces. At the same time, the impulse front spreads radially through the septum parallel to its endocardial surfaces, in each region the subendocardium being involved earlier than the deeper layers. The impulse front reaches the auriculo-ventricular junction before the apex and interventricular grooves. After reaching the auriculo-ventricular junction the impulse front continues to spread toward the apex and at the same time toward the anterior and posterior interventricular grooves. On reaching the interventricular grooves the impulse front spreads over the lateral walls of the right and left ventricles. The front coming from the anterior interventricular groove and spreading in two directions meets that from the posterior interventricular groove, also spreading in two directions, about midway on the lateral surfaces of the right and left ventricles. The order of junction of the impulse fronts in the right and left ventricles is not predictable since the terminal net-work in the right heart travels obliquely from endocardium to epicardium while that of the left is nearly perpendicular to the endocardial surface (1), making up for the shorter distance across the thinner right ventri-

cular wall. However, it can be predicted as a general rule that the basal part of the lateral wall of both ventricles is stimulated last since the horizontal circumference of the heart is greatest at the base. Throughout, the subendocardium is stimulated ahead of the subepicardium and the papillary muscles appear to be stimulated ahead of adjacent muscle.

The Laws Governing the Spread of Current in a Volume Conductor. Some consideration must be given to the laws governing the spread of current in a volume conductor in order to appreciate the effect of the electric forces created in the heart upon the electrocardiogram recorded by electrodes located some distance from the heart. Space does not permit more than a superficial handling of this problem; the reader is referred to several excellent summaries for a more detailed and mathematical handling (13, 27, 61, 299, 314, 356, 367).

The simplest case of a volume conductor is that of a large finite electrically homogeneous sphere with the source of electric current at its center. If the source of current be considered a simple small cell which during its cycle of activity shows a wave of depolarization on its surface and then a wave of repolarization, the basic concepts can be expressed in simple terms. In the completely polarized resting state, it can be demonstrated that no potential differences exist in the outside conductor and no currents flow. Assuming the cell surface to be uniformly polarized, all points in the external conductor are at the same potential, namely zero. Some confusion has arisen in considering this problem because no clear distinction was made between the charge on the cell surface and its potential. Thus, according to the laws of spherical condensors it can be shown that the potential in the region outside of the cell and at its outside surface will be zero even though a positive charge is located on the outside of the cell surface (177, 316).

During depolarization, potential differences will be set up and currents will flow in the outside medium. While the actual source of the current is the polarized membrane remaining on parts of the cell surface, the effect can be accurately and more simply expressed by considering that the current is caused by a surface (or a series of surfaces) at the junction between the polarized and depolarized parts of the cell at any moment. The charge of this theoretical surface is oriented so that the positive charge faces the polarized part of the cell and the negative charge faces the depolarized part of the cell, and the intensity of the charge on the postulated surface is assumed to be identical with that on the actual polarized surface of the cell. A single imaginary surface will be adequate for the case where the process of depolarization at the moment is confined to a small region, i.e., that the rate of depolarization of any area is extremely high compared to the rate of propagation of the process to adjacent areas. If, however, the process of depolarization is relatively slower, then more accurate representation is obtained by assuming a series of planes located at the sites where each increment of depolarization is taking place. Thus, if the intensity of charge of the polarized membrane is considered to be N units, there would be N surfaces placed at the junctions where one unit of charge has been lost, each having a charge of one unit oriented as in the simpler case described above. Ordinarily in depolarization

little is lost if it is assumed that a single theoretical surface depicts the facts, since the series of surfaces which more accurately portray the facts would be very close to one another.

During repolarization, however, since repolarization of each region occupies much more time than does its depolarization, the series of theoretical surfaces would be farther apart and more error is introduced by reducing them to a single theoretical surface. Nevertheless this may still be done to simplify the solution provided one realizes the limitations of the method. When this is done, the single theoretical surface appears to have a smaller charge during repolarization than during depolarization. The orientation of the charges on the theoretical surface during repolarization is such that the positive charge faces the part of the cell already polarized and the negative charge faces that part still depolarized.

With the consideration of such imaginary surfaces, the potential of any point P in the outside medium would have the sign of the charge of the theoretical surface facing it. The magnitude of the potential at P would be a function of the intensity of the charge on the theoretical surface and of the solid angle subtended from the point P to the edge bounding the theoretical surface. Thus if the point P faced the side of the theoretical surface having a positive charge it would have a positive potential. And the intensity of the potential of the point P would increase 1, the closer it was to the theoretical surface; 2, the larger the rim of the theoretical surface; 3, the greater the intensity of charge of the theoretical surface; and 4, for points at equal distance from the center of the theoretical surface, the smaller the angle between the line connecting the point P to the center of the surface and the line from the center of the theoretical surface drawn perpendicular to it. In this fashion the lines of potential differences and the lines of current flow in the external field can be mapped out and the deflection produced in a galvanometer connected to any two points in the electric field predicted.

With some modifications these rules will apply also to the case where the size of the conducting medium is not so large and where it is not spherical.

This concept of theoretical surfaces serves to link the membrane theory with the facts observed in the case of the heart within the body. It explains the finding of a diphasic deflection with brusque transition during activation obtained when an electrode placed on the surface of the auricle is connected to one placed at a distance in the conducting medium. It is an improvement of Lewis' concept of "limited potential differences" (198) as developed by Craib (61) and most clearly enunciated by Wilson (367).

This view of a theoretical surface can also be applied to the case of injury currents. The theoretical surface of the resting injured cell would be at the junction between the injured and uninjured parts. The charge on the theoretical surface would be of a magnitude equal to the intensity of the charge on the uninjured region minus that on the injured part of the cell. The positive charge of the theoretical surface would face the uninjured part of the cell and the negative charge the injured part of the cell. During activity, if the injured part of the cell had retained a polarized state, the theoretical surface would remain but have its charge reversed and of different magnitude. During stimulation the lack

of response of the injured part of the cell, when such unresponsiveness existed, would alter the moment to moment orientation of the theoretical surface which would have existed if no injury were present; and during repolarization the effects of injury in retarding the repolarization of the injured part of the cell would also modify the moment to moment orientation of the theoretical surface which would have existed if no injury were present.

This concept of a theoretical surface will explain the findings 1, that an injury current is obtained when an electrode is placed on an injured part of the heart and connected with an electrode distant from the heart; 2, that this current can be reversed by moving the electrode on the heart to an uninjured region adjacent to the injury (80, 81, 89, 94, 156, 369), and 3, that this current decreases in magnitude rapidly when the electrode on the heart is moved farther away from the site of injury since the angle subtended by the postulated surface diminishes rapidly.

The theoretical surface concept makes it simpler to follow the events during the heart cycle and during injury since the charge giving rise to the electric field can be considered as located at the junction between depolarized and polarized parts of the cell, that is, at the junction between activated and restituted parts of the cell, and the junction between injured and uninjured parts of the cell.

The problem of following events during activation and restitution in the case of a complex cell like the auricles or ventricles can be simplified by this theoretical surface concept. Take the case of the auricles. Ordinarily the impulse spreads from a small area, from the sinus node, over the walls of the auricles caudad, ventrad and to the left, ignoring minor variations. The theoretical plane is therefore at right angles to this spread coinciding with the wave front, with the positive charge ahead of the negative. As it travels over the right auricle the length of the rim of this theoretical plane increases. Actually the wave front is a double ring, the inside ring on the subendocardial and the outer on the subepicardial aspect of the auricle and the theoretical surface being the surface between these rings approximately normal to the endocardial surface. Since it is assumed that the intensity of the polarized surface of the auricular syncytium is uniform, the increasing length of the rim of the theoretical surface will cause the potential differences to wax without much change in direction. As the impulse spreads over the left auricle and septum, the length of the rim will tend to decrease, causing the potential differences to wane. During this time, however, there will be a shift in the general direction of the impulse spread to the left rather than caudad. On the assumption that the topography of repolarization in the auricles follows closely that of the depolarization, the theoretical surface for repolarization will have the same characteristics as that during depolarization but 1, the charge will be reversed, and 2, the theoretical surface will have a less intense charge.

Take the case of the ventricles. As outlined earlier, two theoretical surfaces would start in the interventricular subendocardial surfaces and meet and fuse. During this time no potential differences would appear unless one of these impulses started ahead of the other, and even then the potential differences

created would not be large because the two theoretical surfaces would have oppositely directed charges. Apparently, under ordinary circumstances the stimulation via the left bundle precedes that via the right, making the theoretical plane entering the septum from the left larger than that from the right. After the two theoretical surfaces meet the further spread of the impulse would be radial in the septum, and an expanding narrow cylindrical theoretical surface would be present for a time. No potential differences would arise (except for the small differences explained above) until the cylinder front reached the auriculo-ventricular connective tissue partition. From that time on the narrow cylinder will be open at the top and the part facing toward the apex will therefore be unbalanced. Since it can be shown that regardless of how complex the theoretical surface actually is, it in turn is equivalent to the smallest surface encompassing its rim, such a new theoretical surface can be considered as representing the conditions at this time. In short, the potential field will be determined by this imagined secondary surface at the top of the septum, the positive charge of which faces the apex and the negative the auricles. This surface will expand until the entire junction of the septum with the auriculo-ventricular partition has been depolarized. The intensity of the potential differences will therefore grow. The field will be such as to have its positive potentials caudad, ventrad and to the left, that is, in the same general direction as in the field created during auricular stimulation.

As the impulse reaches and spreads through the lateral walls of the ventricles, some time before the whole base of the septum has been stimulated, the situation becomes more complex. The details have been briefly described above. Suffice it to say that the four wave fronts of impulse spread tend to neutralize one another except that any theoretical surface present on that part of the impulse front which moves toward the apex will be unopposed by any oppositely charged theoretical surface as soon as the wave front reaches the auriculo-ventricular junction. Should the base of the heart be completely activated before the apex then the field created would continue in general to be maintained in the direction found during stimulation of the septum, and will wax and wane. If, however, portions of the base of the heart are activated after the apex is completely stimulated, as is commonly the case, then the field will be markedly altered in its direction toward the end of the depolarization phase.

Injury of the heart situated in a conducting medium alters the electrical deflections obtained. During depolarization of the ventricles, the absence of response of the injured area leaves unbalanced certain theoretical surfaces which ordinarily would be counterbalanced by the response of the injured area, and so the QRS complex is altered. This is especially prominent when a precordial electrode is used over an injured area of the anterior surface of the heart. This alteration of QRS will persist when the area becomes necrotic or is replaced by scar tissue. In reality a new "window" is opened in the wall of the ventricle. Hence QRS changes as a result of injury occur equally readily whether the region is alive and injured, or dead, necrotic or scarred.

Both during the depolarized and repolarized state, the differences in polari-

zation between the injured and uninjured parts of the ventricles will set up a new theoretical surface giving rise to the injury current of rest and/or of activity. Its orientation will determine the potential differences in the field and so will determine the S-T deviation of injury. During repolarization the lag and protraction (or absence) of the repolarization process in the injured area will alter the characteristics of the theoretical surface at this time and set up the T wave alterations of injury.

Vector analysis. In a large finite spherical homogeneous conducting medium with the heart in the center, the electric forces could be resolved during the heart cycle into a series of successive vectors which alter in magnitude and direction. This is so since the electric forces have quantity, sign and direction, and can be added and resolved according to the law of vectors. It is in such a situation that an equilateral triangle can be created to calculate the manifest potential vectors in the frontal plane. In this situation also a tetrahedron composed of equilateral triangles, or a cube, can be created to calculate the actual three dimensional vectors.

Unfortunately the actual situation in the body is such that the attempts to determine the true vectors of the heart are extremely crude and give only a rough approximation of the momentary forces which exist in the heart (163). This is true for a number of reasons: 1, the torso of the body which constitutes the conducting medium is not spherical; 2, it is not a large field compared to the size of the heart (this can be appreciated by comparing the size of the fist to that of the torso); 3, the heart is not in the center but is closer to the sternum than the back, closer to the arms than to the legs and closer to the left than the right arm; 4, the limb leads used in obtaining the vectors in the frontal plane actually do not form an equilateral triangle (163), and 5, the body field is not composed of a homogeneous conducting medium (85, 167, 169, 173, 200), the lungs, for example, being poor conductors compared to the muscles of the back (169, 200). The view held by Wilson's laboratory that the electrical conductivity of the lungs is little different from other tissues is based on impedance measurements with extremely high frequency of alternating current (176). Such high frequency measurements are not applicable, since in electrocardiography we are concerned with fluctuating currents of much lower frequency. Actual experiments with insulation of various surfaces of the heart show clearly that the lungs are much poorer electrical conductors than blood, liver or muscle (169, 200). This gives certain parts of the heart a greater influence over others in creating the electric field of the body.

These discrepancies between conditions in the body and those postulated as ideal conditions for vector analysis are significant enough so that vector analysis such as has been attempted can give only a crude approach to the actual vectors existing during the heart cycle. Of course a rough idea of the direction and possible magnitude of the vectors can be obtained, but the approximation to the actual vectors remains to be established.

Another factor of importance is that the heart is so large that when electrodes are placed on certain parts of the body surface, as over the precordium, equal areas of the heart do not subtend the same angle at the electrode, consequently

they have unequal electric effects. Reduction of the electric forces to a single vector is inadequate since the part of the heart beneath the electrode plays a dominant rôle while more distant parts play a lesser rôle because the contribution of a given area of a polarized membrane to the potential is a function of the distance from the electrode and of the electrical conductivity of the intervening tissues. In this sense, and in this sense only, can the precordial electrode be considered a semidirect electrode. Under these circumstances it is difficult to see how a tetrahedron or cube can give more than a rough approximation of the true three dimensional vector. Attempts to get such stereovectograms are thus only rough approximations.

The construction of vectograms and stereovectograms (9, 27, 53, 54, 70, 74, 79, 135, 137, 149, 150, 151, 152, 154, 302, 303, 304, 305, 319, 320, 321,322,323, 328, 329, 351, 380, 381, 382, 383), the time course respectively of the vector in the frontal plane and in three dimensions, may give a crude picture of the summation of the electric forces existing in the heart from moment to moment during the heart cycle, and can help to approximate the character of the forces in various circumstances. On a purely empiric basis they may prove to be of diagnostic value in various circumstances, but they have not yet been found to serve this purpose. Even if this were accomplished, it cannot be considered that they actually depict the true summation of the electric forces of the heart cycle except in a distorted fashion.

The attempt to summate the vectors during depolarization, during repolarization and during the entire ventricular electrical systole, the resultant vectors being known respectively as the QRS vector, the T (or more properly the S-T-T) vector and the ventricular gradient, (11, 12, 15, 16, 17, 18, 19, 20, 97, 158, 212, 213, 354, 368) may also prove ultimately to be of practical value. However, these summations introduce other errors which may become misleading. For example, the QRS vector of the frontal plane may give an identical value in the case of a deep Q_2 and Q_3 (i.e., a deep Q in leads 2 and 3) as in the case of a deep S_2 and S_3 and yet the frontal vectogram (i.e., a graph of the instantaneous vectors throughout the heart cycle) will be entirely different under these two circumstances. Again the S-T-T vector caused by an inverted coronary T_2 and T_3 and an elevated S-T_2 and S-T_3 may be identical with that obtained when T_2 and T_3 are upright and S-T_2 and S-T_3 are depressed, and yet the frontal vectogram will be entirely different in the two cases. Such summated vectors may thus conceal significant differences and the ventricular gradient derived from them may also be non-revealing. Of course, it is true that the measurements of the ventricular gradient can help to estimate whether a change in the S-T segment and T wave are due to positional alterations of the heart on the one hand, or to injury, heart strain or some influence selectively affecting the repolarization process on the other. In our experience, however, examination of the contour of the electrocardiogram in the limb leads and chest leads will serve this function as well if not better (164). The latter method will in fact avoid the implication that the solution is a mathematical one and therefore necessarily the correct one.

It can be seen from the discussion above that there are two opposing viewpoints.

One overlooks individual variability and seeks a simple mathematical solution. The other is perturbed by the inaccuracies of the mathematical approach and fears that by following it important variables may be neglected. There is merit in both approaches provided zealots do not overestimate or underestimate the relative merits of the opposing viewpoints. Considerable work remains to be done in defining the limits of both points of view. It is as much an error to deny the possibility that the true electric forces of the heart can be reconstructed from the electrocardiogram as it is to affirm that at the present time the estimated vector is adequate, fairly accurate and sufficiently informative for ordinary use.

So-called unipolar leads. Another experimental approach remains to be discussed and that is the attempt to obtain the actual potentials of single points in the conducting medium (25, 52, 55, 111, 112, 113, 114, 115, 129, 179, 219, 220, 355, 356). For practical purposes there is little to be gained from these attempts but the problem of a suitable method to accomplish this is intriguing. If one were dealing with an infinite sphere of a homogeneous conducting medium the potentials of single points might be measurable but with the situation as it exists in the human body this is not strictly possible although approximations may be obtained. The use of the central terminal connected to the three extremities through 5000 ohms each has been widely discussed (cf. 356). Even the proponents who introduced this method recognize that while the central terminal may show less fluctuation in potential than any single distant electrode, it need not be at constant zero potential (317, 356). For example, it can happen that all three limb electrodes may some time during the heart cycle be in the negative field while the positive field is concentrated around the anterior precordium or vice versa. Under these circumstances the central terminal will also have a negative potential (or a positive one) greater than that of one or more electrodes to which it is connected. During this time it will be a "worse" contact for obtaining the true potential of the exploring electrode than some other single electrode. Goldberger's so-called augmented unipolar lead (111, 112, 113, 114, 115) fails on theoretical grounds to meet even the requirements set up for the Wilson terminal, especially when it is connected to only two of the three limb electrodes. Other attempts to obtain the true potential of the exploring electrode have been made. One procedure is to immerse the body in a metal lined tub filled with a strong salt solution (51, 52, 78, 219), the patient breathing by way of a tube passed up through the bath. The entire metal lining of the tub is used as the indifferent null electrode. This is probably as close to a null electrode as one can obtain, but is not too practical. A variant of this procedure is to employ a broad metal covering of a large part of the body which should approach a null electrode more nearly than the central terminal of Wilson (179, 220). However, it would appear that an absolute null electrode has not yet been perfected.

Great care must likewise be employed in using precordial leads to determine the time of arrival of the impulse beneath the chest electrode (356, 357, 364) since distant effects obscure the so-called intrinsic deflection especially when the electrode is not placed over the region of solid dullness.

The Influence Exerted by the Characteristics of the Body as a Conductor of Electri-

city. In the last section it was pointed out that the body differs from the ideal spherical conducting medium of large finite size and that on this account variables are introduced which make accurate calculation of the vectors impossible. These characteristics of the body as a conducting medium also affect the character of the electrocardiogram. Their variation in different individuals and in disease affect electrocardiographic contour. Some of these qualities have unimportant, others more striking effects.

Persons vary in the relative diameters of their torsos. The body field of a fat person is different from that of a thin individual. It is doubtful, however, that these changes in body shape have any significant influence on the electrocardiogram except as they are accompanied by rotation of the heart.

More significant and commonly overlooked is the effect of variations in the electricial conductivity of the medium adjacent to the heart. For example, in emphysema where the lungs, abnormally distended by air, encroach upon the heart and increase the insulation around it, the potential differences set up in the external field tend to be smaller and current flow would largely be confined to the heart and pericardium. As a result, the amplitude of the deflections in the electrocardiogram tends to become smaller. Likewise pericardial or pleural effusions lying close to the heart are good electrical conductors and tend to shunt the external field which is of higher resistance, thereby decreasing the current flow through the body field. Hence, here too, the amplitude of the deflections of the electrocardiogram decreases. The presence of focal pleural effusions or of consolidations in the lungs near the heart tend to produce changes in the electric field of the body and to modify the electrocardiographic contour. In short, alterations of the character of the electrical conductors adjacent to the heart modify the field created and thereby alter the recorded electrical deflections (163). Such changes could affect the magnitude of the vectors reconstructed from the electrocardiogram, without affecting the orientation of the electric forces within the heart.

The Variations Introduced by Modification of the Anatomy and Physiology of the Heart. Variations in the anatomy and physiology of the heart are far more important than variations in the character of the body field in determining electrocardiographic contour.

1. Variations in the normal heart. The recent experience of the war when electrocardiograms on large numbers of normal persons were taken has clearly shown the wide range of variation of the "normal" electrocardiogram caused by differences between individuals. Considerable variation is also found in the same individual in the course of ageing. These variations with age determine the different normal limits of the electrocardiogram in infants, children and adults. The details of the anatomic and physiologic differences of normal hearts in the various age groups will not be considered here (cf. 164, 172, 282). They depend, as do the variations in normal contour in any age group, on 1, the shape and size of the heart; 2, the position of its various axes in relation to the body, and 3, variations in the conduction paths for the impulse and in the location of the impulse origin.

The shape of the heart and the relative size of its various chambers show some variation. This is particularly striking when one compares the heart of an infant with that of an adult. But even in the adult there is a considerable range in shape of the heart. Take for example the heart in which the diameter from base to apex is longer than the other two diameters and compare it with a heart in which the reverse is true. In the first case, assuming uniform rate of impulse spread, the arrival of the impulse at the apex of the septum will be delayed as compared with its arrival near the base of the anterior and posterior interventricular grooves. In the second case, the impulse will reach the apex sooner than the basal parts of the lateral walls of the ventricles at the interventricular grooves. Obviously the topography of the spread of the impulse will be sufficiently different in the two cases to vary the QRS configuration. One could write as extensive a treatise on the effects of such and other differences in the shape of the heart on the QRS vector as Ashman has on the effect of normal changes in the heart's position (97).

The subject of variations in the normal position of the heart has been covered extensively elsewhere (5, 6, 14, 97, 357) and need not be recapitulated here. It is apparent that this is of importance in accounting for the wide range in contour of the various deflections of the normal electrocardiogram.

A similar effect can be demonstrated for variations in the anatomy and function of the specialized muscle of the heart from which the impulse originates and over which it spreads. Changes in the P wave may come about not only because of altered shape of the auricles or alterations in the position of their axes relative to the body but also because the pacemaker may be located in the body or tail of the sinus node instead of the head. Similar P wave contour changes may occur when the site at which the impulse leaves the sinus node is different from the usual point of exit, even though the point of its origin has not shifted.

The anatomy of the specialized conducting system in the ventricles shows a normal range of variation from one heart to another, and these variations together with the variations in the speed of impulse transmission through the various elements of the bundle branch system and the associated terminal Purkinje net-work participate in the production of variations in the configuration of the ventricular deflections in the electrocardiogram.

Not enough attention has been paid to the factors just discussed, namely, the variations in the normal anatomy and physiology of impulse origin and spread, the normal variations in size and shape of the heart and the normal variations in the location and properties of the electric conductors adjacent to the heart in comparison to the extensive consideration given to the normal variations in the location of the axes of the heart. As a result, the latter have been given undue emphasis and some workers have resorted to nebulous terms like "electrocardiographic vertical position of the heart" to explain certain patterns of the electrocardiogram not revealed by fluoroscopic or x-ray examinations of the heart.

2. Variations in the abnormal heart. Changes in shape, size and position of the heart play an important rôle in determining the electrocardiographic contour in the abnormal heart (184). Dilatation of the various chambers, displacements of

the entire heart and rotation on the various axes of the heart are common in disease and play a significant rôle in determining the appearance of the electrocardiogram. These factors have only recently been given adequate consideration in evaluating abnormalities of the electrocardiogram.

Heart disease also produces electrocardiographic alterations by causing hypertrophy of the various chambers leading thereby to heart strain patterns. These come about because of the prolongation and altered manner of impulse spread in the hypertrophied chamber which prolongs the time of depolarization to a slight degree, alters the balance of electric forces during this time and tends to make the pattern of repolarization follow more nearly the pattern of depolarization in time and space than is normally the case.

Other pathological changes in electrocardiographic contour are caused by regions of ischemia and injury. By leading to unresponsive areas, to injury currents of rest, of activity and/or of repolarization, they are responsible for many modifications of the electrocardiogram. Sometimes these are sufficiently striking so as to constitute specific patterns easily recognizable in the electrocardiogram.

Disease may modify the electrocardiogram by altering the speed of repolarization in the whole heart or in different parts of the heart. Finally, disease leads to abnormalities in electrocardiographic contour by displacing the location of impulse origin and by producing delays or blockage in various parts of the path over which the impulse spreads.

Drugs operate in fashions similar to disease in causing electrocardiographic alterations.

It is the task of the electrocardiographer on the basis of experience and some knowledge of the theory of the genesis of the electrocardiogram to interpret properly the abnormalities in its contour. How this is accomplished has been considered at length elsewhere (164). A major aim of these efforts to unravel the genesis of the electric forces of the heart must be the advancement of the ability to properly evaluate the electrocardiogram for clinical utility.

SUMMARY

The present review concerns itself with the theoretical aspects of the electrocardiogram. Current theories and findings are critically evaluated from the author's viewpoint.

The subject is developed under the following four headings:

a. The genesis and typical time course of the electric states of the heart during activity and injury.

b. The laws governing the spread of current within a volume conductor.

c. The influence exerted by the characteristics of the body as a conductor of electricity.

d. The modifications introduced by variations of the anatomy and physiology of the heart.

a. The genesis of the electric states of the heart is discussed in terms of the classical hypothesis of polarization, depolarization and repolarization of the cell

surface corresponding to rest, activity and recovery, respectively. Modern refinements of these concepts do not significantly modify the classical concepts of the production of the main deflections recorded in the electrocardiogram. Alterations of the normal record caused by regional injuries to the heart muscle are explained and classified according to the mode of action of injury; injuries may be such as to produce changes during rest, activity or the recovery phase, or lead to a combination of these. The present state of the problem of neurogenic versus myogenic impulse spread and of the syncytial character of heart muscle is discussed. The view is expressed and defended that, some recent evidence to the contrary, myogenic transmission accounts for the spread of the impulse, and that, on the basis of their syncytial character, the auricles and the ventricles each acts electrically as a single cell. A brief detailed topographical and chronological account is presented of the spread of the impulse as the basis on which the typical time sequence of the deflections seen in the electrocardiogram can be understood.

b. A discussion follows of the modification introduced into the recording of the electric states of the heart produced by the circumstance that the recording electrodes are placed at a distance from the source of current. These modifications are considered in terms of theoretical charged surfaces interposed between polarized and depolarized regions of the heart which alter in position and direction of charge as the sources and sinks of current shift during the heart cycle. The influence of injury upon such theoretical charged surfaces is also considered. Next, the attempt to summate the electric forces created during the heart cycle in terms of vector analysis is presented. Its main limitation is seen in the fact that the initial assumptions from which the analysis proceeds are, as yet, crude oversimplifications which must lead to a distorted picture of events in the heart.

c. Aside from the distortions generally introduced by considering the body as a large sphere of homogeneous conducting properties with the heart at its center, individual distortion may be introduced by pathological variation of the conducting media next to the heart.

d. Considerably more importance is attached to the effect of modifications in the anatomy and physiology of the heart itself in altering the contour and time pattern of the normal electrocardiogram. As examples of such modifications of the normal heart are discussed differences in size and shape between the infantile and the adult heart, the effect of variations in the relative length of the two main heart diameters from one heart to another, and the well worked out effect of variations in the position of the main axes of the heart with respect to the body. Of even more striking effect than these normal variations are, of course, pathological variations in the anatomy and physiology of the heart which are encountered in disease.

REFERENCES

(Essentially the significant literature since 1930)

(1) Abramson, D. I. and J. C. Cardwell. A new anatomic basis for the spread of the impulse in the mammalian ventricle. Am. J. Physiol. 109: 1, 1934.

(2) ABRAMSON, D. I. AND K. JOCHIM. The spread of the impulse in the mammalian ventricle. Am. J. Physiol. 120: 635, 1937.

(3) ABRAMSON, D. I., L. N. KATZ, S. MARGOLIN AND S. LOURIE. Variations in the electrocardiographic form of experimental ventricular ectopic beats induced in the monkey and dog. Am. Heart J. 13: 217, 1937.

(4) ABRAMSON, D. I. AND S. MARGOLIN. A Purkinje conduction network in the myocardium of the mammalian ventricles. J. Anat. 70: 250, 1936.

(5) ABRAMSON, D. I. AND J. WEINSTEIN. A basis for the analysis of variations in the form of electrocardiographic curves resulting from experimental premature contractions. Am. J. Physiol. 115: 569, 1936.

(6) ACKERMAN, W. AND L. N. KATZ. Reversal in direction of the QRS complex of experimental right bundle branch block with change in the heart's position. Am. Heart J. 8: 490, 1933.

(7) AMUCHASTEGUI, S. R. O., O. ORIAS AND A. S. SEGURA. Variaciones de los complejos inicial y final de electrogramas miocardicos segun la dirección de propagación del estímulo. Rev. Soc. Argent. Biol. 18: 138, 1942.

(8) ARMSTRONG, P. B. Functional reactions in the embryonic heart accompanying the ingrowth and development of the vagus innervation. J. Exper. Zool. 58: 43, 1931.

(9) ARRIGHI, F. P. El eje eléctrico del corazón en el espacio. Prensa Med. Argent. 26: 253, 1939.

(10) ASHMAN, R. An outline of electrocardiography. New International Clinics 1: 193, 1939.

(11) ASHMAN, R. The normal human ventricular gradient. IV. The relationship between the magnitudes, AQRS and G, and deviations of the RS-T segment. Am. Heart J. 26: 495, 1943.

(11a) ASHMAN, R. Heart. Ann. Rev. Physiol. 6: 319, 1944.

(12) ASHMAN, R. A statistical study of the ventricular gradient and of the QRS complex of the electrocardiogram. Arch. Inst. Cardiología México 15: 266, 1945.

(13) ASHMAN, R. Nociones fundamentales de electrocardiografía. Rev. Argent. Cardiol. 12: 220, 1945.

(14) ASHMAN, R. Estimation of heart position from the QRS complex of the electrocardiogram. Arch. Inst. Cardiología México 16: 139, 1946.

(15) ASHMAN, R., E. BYER AND R. H. BAYLEY. The normal human ventricular gradient. I. Factors which affect its direction and its relation to the mean QRS axis. Am. Heart J. 25: 16, 1943.

(16) ASHMAN, R. AND E. BYER. The normal human ventricular gradient. II. Factors which affect its manifest area and its relationship to the manifest area of the QRS complex. Am. Heart J. 25: 36, 1943.

(17) ASHMAN, R., F. P. FERGUSON, A. I. GREMILLION AND E. BYER. The effect of cycle-length changes upon the form and amplitude of the T deflection of the electrocardiogram. Am. J. Physiol. 143: 453, 1945.

(18) ASHMAN, R., F. P. FERGUSON, A. I. GREMILLION AND E. BYER. The normal human ventricular gradient. Am. Heart J. 29: 697, 1945.

(19) ASHMAN, R, F. P. FERGUSON, A. I. GREMILLION AND E. BYER. Effect of cardiac cycle-length upon magnitude of ventricular gradient. Proc. Soc. Exper. Biol. and Med. 59: 47, 1945.

(20) ASHMAN, R., M. GARDBERG AND E. BYER. The normal human ventricular gradient. III. The relation between the anatomic and electrical axes. Am. Heart J. 26: 473, 1943.

(21) ASHMAN, R. AND R. HAFKESBRING. Unidirectional block in heart muscle. Am. J. Physiol. 91: 65, 1929.

(22) ASHMAN, R., W. S. WILDE AND C. E. DRAWE. The validity of the electrical doublet theory of the cardiac action current. Am. J. Physiol. 128: 547, 1940.

(23) ASHMAN, R. AND N. C. WOODY. Monophasic action currents from the uninjured turtle ventricle. Proc. Soc. Exper. Biol. and Med. 42: 17, 1939.

(24) Barker, P. S., A. G. Macleod and J. Alexander. The excitatory process observed in the exposed human heart. Am. Heart J. 5: 720, 1930.

(25) Barker, P. S., F. N. Wilson and F. D. Johnston. Electrocardiograms that record the potential variations produced by the heart beat at a single point. J. Clin. Investigation 12: 969, 1933.

(26) Bayley, R. H. The potential produced by cardiac muscle. A general and a particular solution. Proc. Soc. Exper. Biol. and Med. 42: 699, 1939.

(27) Bayley, R. H. On certain applications of modern electrocardiographic theory to the interpretation of electrocardiograms which indicate myocardial disease. Am. Heart J. 26: 769, 1943.

(28) Bayley, R. H. The electrocardiographic effects of injury at the endocardial surface of the left ventricle. Am. Heart J. 31: 677, 1946.

(29) Bayley, R. H. and J. S. La Due. Electrocardiographic changes of impending infarction. Am. Heart. J. 28: 54, 1944.

(30) Bayley, R. H. and J. S. La Due. Differentiation of the electrocardiographic changes produced in the dog by prolonged temporary occlusion of a coronary artery from those produced by postoperative pericarditis. Am. Heart J. 28: 233, 1944.

(31) Bayley, R. H., J. S. La Due and D. J. York. Electrocardiographic changes (local ventricular ischemia and injury) produced in the dog by temporary occlusion of a coronary artery, showing a new stage in the evolution of myocardial infarction. Am. Heart J. 27: 164, 1944.

(32) Bayley, R. H., J. S. La Due and D. J. York. Further observations on the ischemia-injury pattern produced in the dog by temporary occlusion of a coronary artery. Am. Heart J. 27: 657, 1944.

(33) Benjamin, J. E., H. Landt and L. R. Culver. The body as a volume conductor and its influence on the electrical field of the heart. Am. J. Med. Sci. 195: 759, 1938.

(34) Bishop, G. H. La théorie des circuits locaux permet-elle de prévoir la forme du potentiel d'action? Arch. Intern. Physiol. 45: 273, 1937.

(35) Bishop, G. H. The relation of bioelectric potentials to cell functioning. Ann. Rev. Physiol. 3: 1, 1941.

(36) Bishop, G. H. and A. S. Gilson. Action potentials from skeletal muscles. Am. J. Physiol. 89: 135, 1929.

(37) Blair, D. M. and F. Davies. Observations on the conducting system of the heart. J. Anat. 69: 303, 1935.

(38) Blair, D. M., F. Davies and E. T. B. Francis. Conducting system of marsupial heart. Trans. Roy. Soc. Edinburgh 60: 629, 1943.

(39) Blair, H. A., A. M. Wedd and A. C. Young. The relation of the Q-T interval to the refractory period, the diastolic interval, the duration of contraction and the rate of beating in heart muscle. Am. J. Physiol. 132: 157, 1941.

(40) Bogue, J. Y. The electrocardiogram of the developing chick. J. Exper. Biol. 10: 286, 1933.

(41) Bohning, A., L. N. Katz and R. Langendorf. The distribution of surface potential on the chest in intraventricular block. Am. Heart J. 22: 778, 1941.

(42) Bohning, A., L. N. Katz, R. Langendorf and B. Blumenthal. Intraventricular block, including so-called bundle branch block. Am. J. Med. Sci. 202: 671, 1941.

(43) Bohning, A., L. N. Katz, M. Robinow and G. Gertz. The value and significance of multiple chest leads in man. I. Normal and hypertrophied hearts. Am. Heart J. 18: 25, 1939.

(44) Boyd, L. J. and D. Scherf. The electrocardiogram after mechanical injury of the inner surface of the heart. Bull. N. Y. Med. College 3: 1, 1940.

(45) Bozler, E. The initiation of impulses in cardiac muscle. Am. J. Physiol. 138: 273, 1943.

(46) Brown, W. H. A study of the esophageal lead in clinical electrocardiography. I. Am. Heart. J. 12: 1, 1936.

(47) BROWN, W. H. A study of the esophageal lead in clinical electrocardiography. II. Am. Heart J. **12**: 307, 1936.
(48) BUCHBINDER, W. AND L. N. KATZ. The electrocardiogram in acute experimental distention of the right heart. Am. J. Med. Sci. **187**: 785, 1934.
(49) BÜCHNER, F. Die Deutung des Elektrokardiogramms bei den Durchblutungsstörungen des Herzmuskels. Klin. Wchnschr. **17**: 1713, 1938.
(50) BURGE, W. E., O. S. ORTH, H. W. NEILD, R. KROUSE AND G. C. WICKWIRE. Cause and significance of electronegativity of active living tissue. J. Lab. Clin. Med. **21**: 1162, 1936.
(51) BURGER, R. Über das elektrische Feld des Herzens; über die herzferne Electrode bei der semidirekten Ableitung; über die Nullpotentialelektrode von Wilson; über die Verwendung einer von der gesamten Körperoberfläche ableitenden Elektrode als indifferente Test-Elektrode. Cardiologia **3**: 56, 1939.
(52) BURGER, R. AND F. WUHRMAN. Über das elektrische Feld des Herzens. II. Cardiologia **3**: 139, 1939.
(53) BURGER, R. AND F. WUHRMAN. Über elektrokardiographische Analyse mittles Vektordiagramm und elektrischem Feld; Demonstration an einem Vorderwand-Spitzen-Infarkt mit W-form des Initialkomplexes in den Ableitungen I and II und kleinen Ausschlägen in allen Extremitäten-Ableitungen. Schweiz. med. Wchnschr. **71**: 65, 1941.
(54) BURGER, R. AND F. WUHRMAN. Der Erregungsablauf im Herzen und das elektrische Feld, demonstriert an einem Fall von sogenanntem Wilson-Block. Schweiz. med. Wchnschr. **71**: 261, 1941.
(55) CABRERA, E. AND D. SODI PALLARES. Determinación del eje medio manifiesto por las derivaciones unipolares de los miembros. Arch. del Inst. Cardiologica México **14**: 135, 1945.
(56) CALABRESI, M. AND A. J. GEIGER. Potential changes in injured cardiac muscle. Am. J. Physiol. **137**: 440, 1942.
(57) CANFIELD, R. On the electrical field surrounding doublets and its significance from the standpoint of Einthoven's equations. Heart **14**: 102, 1927.
(58) CARDWELL, J. C. AND D. I. ABRAMSON. The atrioventricular conduction system of the beef heart. Am. J. Anat. **49**: 167, 1931.
(59) COLE, K. S. AND H. J. CURTIS. Electrical impedance of Nitella during activity. J. Gen. Physiol. **22**: 37, 1938.
(59a) COLE, K. S. AND R. F. BAKER. Longitudinal impedance of the squid giant axon. J. Gen. Physiol. **24**: 771, 1940–41.
(60) CORREA, R. M. E. Citología del miocardio y sistema de conducción aurículo-ventricular (en el perro). Rev. Soc. Argent. Biol. **14**: 486, 1938.
(61) CRAIB, W. H. The electrocardiogram. Privy Council, Med. Research Council, Spec. Report, Series no. 147, 1930.
(62) CURTIS, H. J. AND K. S. COLE. Membrane action potentials from the squid giant axon. J. Cell. Comp. Physiol. **15**: 147, 1940.
(63) CURTIS, H. J. AND K. S. COLE. Membrane resting and action potentials of the squid giant axon. Am. J. Physiol. **133**: 254, 1941.
(64) DAVIES, F. The conducting system of the bird's heart. J. Anat. **64**: 129, 1930.
(65) DAVIES, F. Further studies of the conducting system of the bird's heart. J. Anat. **64**: 319, 1930.
(66) DAVIES, F. The conducting system of the monotreme heart. J. Anat. **65**: 339, 1931.
(67) DAVIES, F. The conducting system of the vertebrate heart. Brit. Heart J. **4**: 66, 1942.
(68) DAVIES, F. AND E. T. B. FRANCIS. The heart of the salamander with special reference to the conducting (connecting) system and its bearing on the phylogeny of the conducting systems of mammalian and avian hearts. Phil. Trans. Royal Soc. London Series B., **231**: 99, 1941.

(69) De Boer, S. Researches on the electrocardiogram. Cardiologia 2: 292, 1938.
(70) Decherd, G. M. and A. Ruskin. Momentary atrial electrical axes. I. Normal sinus rhythm. Am. Heart J. 28: 794, 1944.
(71) Decker, M. Über den Anstieg der Aktionsspannungskurve des Ventrikels. Arch. f. Kreislauf. 6: 75, 1940.
(72) Donnan, F. G. The theory of membrane equilibria. Chemical Rev. 1: 73, 1924.
(73) Drury, A. N. and D. W. Mackenzie. The influence of vagal stimulation upon conduction through the branches of the A-V bundle in the dog. J. Physiol. 80: 329, 1934.
(74) Duchosal, P. W. Physiopathologie de l'electrocardiogramme. Helv. Med. Acta. 12: 361, 1945.
(75) Eccles, J. C. and H. E. Hoff. The rhythm of the heart beat. I. Location, action potential and electrical excitability of the pacemaker. Proc. Roy. Soc. B. 115: 307, 1934.
(76) Eccles, J. C. and H. E. Hoff. The rhythm of the heart beat. II. Disturbance of rhythm produced by late premature beats. Proc. Roy. Soc. B. 115: 327, 1934.
(77) Eccles, J. C. and H. E. Hoff. The rhythm of the heart beat. III. Disturbances of rhythm produced by early premature beats. Proc. Roy. Soc. B. 115: 352, 1934.
(78) Eckey, P. and R. Fröhlich. Zur Frage der unipolaren Ableitung des Elektrokardiogramms. Arch. f. Kreislauf. 2: 349, 1938.
(79) Einthoven, W., G. Fahr and A. de Waart. Über die Richtung und die manifeste Grösse der Potentialschwankungen im menschlichen Herzen und über den Einfluss der Herzlage auf die Form des Elektrokardiogramms. Pflüger's Arch. 150: 308, 1913.
(80) Eyster, J. A. E. and W. E. Gilson. The development and contour of cardiac injury potential. Am. J. Physiol. 145: 507, 1946.
(81) Eyster, J. A. E. and H. Goldberg. The relation of injury potentials in heart muscle to other electrical and to mechanical events. Am. J. Physiol. 133: 272, 1941.
(82) Eyster, J. A. E., M. R. Krasno and J. P. Hettwer. Electrical potentials of the heart of the chick embryo. Am. J. Physiol. 120: 173, 1937.
(83) Eyster, J. A. E., M. R. Krasno, C. A. Maaske and M. J. Ulevich. The origin of the R and T potentials from the mammalian heart. Am. J. Physiol. 120: 663, 1937.
(84) Eyster, J. A. E., F. Maresh and M. R. Krasno. The electric field around heart muscle. Am. J. Physiol. 105: 33, 1933.
(85) Eyster, J. A. E., F. Maresh and M. R. Krasno. The nature of the electrical field around the heart. Am. J. Physiol. 106: 574, 1933.
(86) Eyster, J. A. E., F. Maresh and M. R. Krasno. The electrical field around heart and skeletal muscle. Am. J. Physiol. 109: 34, 1934.
(87) Eyster, J. A. E., F. Maresh and M. R. Krasno. The nature of the R wave potentials in the tortoise and frog heart. Am. J. Physiol. 110: 422, 1934.
(88) Eyster, J. A. E. and W. J. Meek. The sequence of fractionate contraction at different surface regions on the right auricle and ventricles of the dog's heart. Am. J. Physiol. 134: 513, 1941.
(89) Eyster, J. A. E. and W. J. Meek. Cardiac injury potentials. Am. J. Physiol. 138: 166, 1942.
(90) Eyster, J. A. E., W. J. Meek and W. E. Gilson. Spread of impulse over the ventricles in the mammalian heart. Am. J. Physiol. 123: 61, 1938.
(91) Eyster, J. A. E., W. J. Meek and H. Goldberg. The relation of the electrical and mechanical events in the dog's heart and the spread of activity in the right auricle and both ventricles. Am. J. Physiol. 129: 354, 1940.
(92) Eyster, J. A. E., W. J. Meek and H. Goldberg. The relation between electrical and mechanical events in the dog's heart. Am. J. Physiol. 131: 760, 1941.

(93) EYSTER, J. A. E., W. J. MEEK, H. GOLDBERG AND H. L. BARTSCH. The sequence of onset of injury potentials on the surface of the dog and turtle ventricle. Am. J. Physiol. **126**: 488, 1939.

(94) EYSTER, J. A. E., W. J. MEEK, H. GOLDBERG AND W. E. GILSON. Potential changes in an injured region of cardiac muscle. Am. J. Physiol. **124**: 717, 1938.

(95) FAHR, G. AND A. WEBER. Über die Ortsbestimmung der Erregung in menschlichen Herzen mit Hilfe der Elektrokardiographie. Deutsch. Arch. f. klin. Med. **117**: 361, 1915.

(96) FEIL, H. S., L. N. KATZ, R. A. MOORE AND R. W. SCOTT. The electrocardiographic changes in myocardial ischemia. Am. Heart J. **6**: 522, 1931.

(97) GARDBERG, M. AND R. ASHMAN. The QRS complex of the electrocardiogram. Arch. Int. Med. **72**: 210, 1943.

(98) GARREY, W. E. AND C. E. KING. Impulse transmission in the sinus venosus of the turtle heart. Am. J. Physiol. **133**: 288, 1941.

(99) GARREY, W. E. AND C. E. KING. Localization of the pacemaker in the sinus venosus of the turtle heart.' Am. J. Physiol. **133**: 288, 1941.

(100) GÉRAUDEL, E. La théorie vestibulaire du mécanisme cardiaque. Presse Méd. **43**: 297, 1935.

(101) GILSON, A. S. Factors determining block of the conducted cardiac impulse. Am. J. Physiol. **110**: 376, 1935.

(102) GILSON, A. S. AND G. H. BISHOP. The interpretation of electrocardiograms obtained from heart strips in a conducting medium. Am. J. Physiol. **109**: 40, 1934.

(103) GILSON, A. S. AND G. H. BISHOP. The effect of remote leads upon the form of the recorded electrocardiogram. Am. J. Physiol. **118**: 743, 1937.

(104) GILSON, W. E., R. E. CONKLIN AND J. A. E. EYSTER. The relation of fractionate contraction at various surface regions of the ventricles of the snapping turtle to the surface potential distribution. Fed. Proc. **1**: 29, 1942.

(105) GLOMSET, D. J. AND R. F. BIRGE. A morphologic study of the cardiac conduction system. IV. The anatomy of the upper part of the ventricular septum in man. Am. Heart J. **29**: 526, 1945.

(106) GLOMSET, D. J. AND A. T. A. GLOMSET. A morphologic study of the cardiac conduction system in ungulates, dog and man. I. The sino-atrial node. Am. Heart J. **20**: 389, 1940.

(107) GLOMSET, D. J. AND A. T. A. GLOMSET. A morphologic study of the cardiac conduction system in ungulates, dog and man. II. The Purkinje system. Am. Heart J. **20**: 677, 1940.

(108) GLOMSET, D. J., A. T. A. GLOMSET, B. D. GLOMSDT AND R. F. BIRGE. A morphologic study of the cardiac condution system. III. Bundle branch block. Am. Heart J. **28**: 348, 1944.

(109) GOLDBERG, H. AND J. A. E. EYSTER. The electrical field on the surface of the active turtle ventricle. Am. J. Physiol. **126**: 503, 1939.

(110) GOLDBERG, H. AND J. A. E. EYSTER. The relation of electrical and mechanical events in the ventricle of the turtle. Am. J. Physiol. **128**: 390, 1940.

(111) GOLDBERGER, E. A simple, indifferent, electrocardiographic electrode of zero potential and a technique of obtaining augmented, unipolar, extremity leads. Am. Heart J. **23**: 483, 1942.

(112) GOLDBERGER, E. The aVl, aVr, and aVf leads; simplification of a standard lead electrocardiographically. Am. Heart J. **24**: 378, 1942.

(113) GOLDBERGER, E. Use and advantages of augmented unipolar extremity leads in electrocardiographic diagnosis of myocardial infarction. N. Y. State J. Med. **43**: 961, 1943

(114) GOLDBERGER, E. Advantages of augmented unipolar extremity leads (aV leads) over the standard leads in clinical electrocardiography. Rev. Argent. Cardiol. **11**: 149, 1944.

(115) GOLDBERGER, E. Studies on unipolar leads. IV. Effects of digitalis. Am. Heart J. **28:** 370, 1944.
(116) GOLDBERGER, E. An interpretation of axis deviation and ventricular hypertrophy. Am. Heart J. **28:** 621, 1944.
(117) GOLDBERGER, E. The validity of the Einthoven triangle hypothesis. Am. Heart J. **29:** 369, 1945.
(118) GOLDBERGER, E. The differentiation of normal from abnormal Q waves. Am. Heart J. **30:** 341, 1945.
(119) GOLDBERGER, E. Significance of downward T waves in precordial leads of normal children. Am. J. Dis. Children **71:** 618, 1946.
(120) GOLDENBERG, M. AND C. J. ROTHBERGER. Über das Elektrogramm der spezifischen Herzmuskulature. Pflüger's Arch. **237:** 295, 1936.
(121) GROEDEL, F. M. How far does the situation of the indifferent electrode influence the electrocardiographic unipolar chest leads. Cardiologia **3:** 23, 1938.
(122) GROEDEL, F. M. Topography and time of appearance of the action-potential of the heart on the anterior and posterior chest wall in young healthy persons. Cardiologia **4:** 1, 1940.
(123) GROEDEL, F. M. AND P. R. BORCHARDT. Extrasystoles in man produced by direct mechanical irritation of the left or right ventricle. Exper. Med. and Surg. **4:** 145, 1946.
(124) GROEDEL, F. M., B. KISCH AND P. REICHERT. Changes in the standard electrocardiogram and the chest leads during the first stages of life. Cardiologia **6:** 1, 1942.
(125) GROEDEL, F. M. AND E. KOCH. Tierversuche zur Frage des Partial-Elektrokardiogramms der rechten und der linken Herzkammer. Ztschr. f. Kreislauf. **25:** 794, 1933.
(126) HAAGER, B. AND A. WEBER. Klinische und experimentelle Untersuchung über das Elektrokardiogramm; über monophasiche Ableitung vom Tierherzen. Ztschr. f. klin. Med. **131:** 136, 1936.
(127) HARRIS, A. S. Spread of excitation in turtle heart as measured by remote and exploring electrodes and two contiguous electrodes. Am. J. Physiol. **129:** 376,1940.
(128) HARRIS, A. S. The spread of excitation in turtle, dog, cat and monkey ventricles. Am. J. Physiol. **134:** 319, 1941.
(129) HECHT, H. H. The influence of the indifferent electrode upon the precordial electrocardiogram; I. The normal electrocardiogram. Am. Heart J. **24:** 529, 1942.
(130) HECHT, H. H. Potential variations of the right auricular and ventricular cavities in man. Am. Heart J. **32:** 39, 1946.
(131) HELM, J. D., JR., G. H. HELM AND C. C. WOLFERTH. The distribution of potential of ventricular origin below the diaphragm and in the esophagus. Am. Heart J. **27:** 755, 1944.
(132) HERVE L. AND F. HUIDOBRO. Posición del corazón y forma de extrasístoles. Anal. Acad. Biol. Univ. Chile **3:** 43, 1939.
(133) HESS, W. Modellversuche über den Verlauf der Potentiallinien des Herzens im transversalen Brustkorbquerschnitt. Ztschr. f. Kreislauff. **27:** 433, 1935.
(134) HILL, A. Analysis of the normal QRS deflection. Lancet **2:** 1110, 1938.
(136) HILL, I. G. W., F. D. JOHNSTON AND F. N. WILSON. The form of the electrocardiogram in experimental myocardial infarction. Am. Heart J. **16:** 309, 1938.
(137) HILL, R. F. The construction of the cardiac vector. Am. Heart J. **32:** 72, 1946.
(138) HOFF, E. C., T. C. KRAMER, D. DUBOIS AND B. M. PATTEN. The development of the electrocardiogram of the embryonic heart. Am. Heart J. **17:** 470, 1939.
(139) HOFF, H. E. AND L. H. NAHUM. The supernormal period in the mammalian ventricle. Am. J. Physiol. **124:** 591, 1938.
(140) HOFF, H. E. AND L. H. NAHUM. Influence of local applications of potassium chloride on action current of the mammalian heart. Proc. Soc. Exper. Biol. and Med. **45:** 263, 1940.

(141) Hoff, H. E. and L. H. Nahum. The factors determining the direction of the T wave; the effect of heat and cold upon the dextro- and levo-cardiogram. Am. J. Physiol. **131**: 700, 1941.

(142) Hoff, H. E. and L. H. Nahum. A comparison of the configuration in the electrocardiogram of endocardial and epicardial extrasystoles. Am. J. Physiol. **140**: 148, 1943.

(143) Hoff, H. E. and L. H. Nahum. The electrocardiographic localization of myocardial infarcts by injury currents and ventricular extrasystoles. Am. J. Physiol. **143**: 723, 1945.

(144) Hoff, H. E., L. H. Nahum and W. Kaufman. The nature of leads I and III of the electrocardiogram. Am. J. Physiol. **134**: 390, 1941.

(145) Hoff, H. E., L. H. Nahum and W. Kaufman. The nature of QI and QIII. Am. J. Physiol. **135**: 752, 1942.

(146) Hoff, H. E., L. H. Nahum and W. Kaufman. Distribution in leads I, II, and III of potentials applied to the surface of the heart. Am. J. Physiol. **138**: 644, 1943.

(147) Hoff, H. E., L. H. Nahum and B. Kisch. Influence of right and left ventricles on the electrocardiogram. Am. J. Physiol. **131**: 687, 1941.

(148) Hogg, M. B., C. M. Goss and K. S. Cole. Potentials in embryo rat heart muscle cultures. Proc. Soc. Exper. Biol. and Med. **32**: 304, 1935.

(149) Hollmann, H. E. and W. Hollmann. Das Einthoven'sche Dreiecksschema als Grundlage neuer elektrokardiographischer Registriermethoden. Ztschr. f. klin. Med. **134**: 732, 1938.

(150) Hollmann, H. E. and W. Hollmann. Das Einthovensche Dreiecksschema im Vergleich zu anderen Ableitungsschemen. Arch. f. Kreislauff. **3**: 191, 1938.

(151) Hollmann, H. E. and E. Guckes. Das Triogram und seine klinische Bedeutung. Arch. f. Kreislauff. **4**: 69, 1939.

(152) Hollmann, W. and H. E. Hollmann. Neue Elektrokardiographische Untersuchungsmethoden. Ztschr. f. Kreislauff. **29**: 546, 1937.

(153) Holzmann, M. Neuere praktische Ergebnisse der Elektrokardiographie. Schweiz. med. Wchnschr. **69**: 289, 1939.

(154) Howard, F. H. A method for the construction of the vectocardiogram from the Einthoven electrocardiogram. Am. Heart J. **31**: 191, 1946.

(155) Jervell, A. Über die theoretische Grundlage der Elektrokardiographie und ihre praktische Bedeutung. Acta Med. Scandinav. **78**: 41, 1936.

(156) Jochim, K., L. N. Katz and W. Mayne. The monophasic electrogram obtained from the mammalian heart. Am. J. Physiol. **111**: 177, 1935.

(157) Johnston, F. D., I. G. W. Hill and F. N. Wilson. The form of the electrocardiogram in experimental myocardial infarction. Am. Heart J. **10**: 889, 1935.

(158) Johnston, F. D. and F. N. Wilson. A vacuum tube integrator for measuring the areas of waves of the electrocardiogram. Proc. Centr. Soc. Clin. Res. **16**: 55, 1943.

(159) Katz, L. N. The asynchronism of right and left ventricular contractions and the independent variations in their duration. Am. J. Physiol. **72**: 655, 1925.

(160) Katz, L. N. Effect on electrocardiogram of opening the thorax and inserting optical manometers into the aorta and pulmonary artery. Proc. Soc. Exper. Biol. and Med. **24**: 652, 1927.

(161) Katz, L. N. The significance of the T wave in the electrogram and electrocardiogram. Physiol. Rev. **8**: 447, 1928.

(162) Katz, L. N. Recent advances in the interpretation of the electrocardiogram. J.A.M.A. **97**: 1364, 1931.

(163) Katz, L. N. Concerning a new concept of the genesis of the electrocardiogram. Am. Heart J. **13**: 17, 1937.

(164) Katz, L. N. Electrocardiography. Lea & Febiger, Philadelphia, 2nd ed., 1946.

(165) KATZ, L. N. AND W. ACKERMAN. The effect of the heart's position on the electrocardiographic appearance of ventricular extrasystoles. J. Clin. Investigation **11:** 1221, 1932.
(166) KATZ, L. N., A. M. GOLDMAN, R. LANGENDORF, L. G. KAPLAN AND S. T. KILLIAN. The diagnostic value of the electrocardiogram based on an analysis of 149 autopsy cases. Am. Heart J. **24:** 627, 1942.
(167) KATZ, L. N., I. GUTMAN AND F. H. OCKO. Alterations of the electrical field produced by changes in the contacts of the heart with the body. Am. J. Physiol. **116:** 302, 1936.
(168) KATZ, L. N., K. JOCHIM AND A. GOLDMAN. The effect of an injured area on the electrical field of the heart based on experiments with models. Am. J. Physiol. **137:** 779, 1942.
(169) KATZ, L. N. AND H. KOREY. The manner in which the electric currents generated by the heart are conducted away. Am. J. Physiol. **111:** 83, 1935.
(170) KATZ, L. N. AND R. LANGDENDORF. The value of chest leads CF_2, CF_4 and CF_5 in the evaluation of heart strain. Arch. Inst. Cardiologica México **15:** 43, 1945.
(171) KATZ, L. N., H. LANDT AND A. BOHNING. The delay in the onset of ejection of the left ventricle in bundle branch block. Am. Heart J. **10:** 681, 1935.
(172) KATZ, L. N. AND M. ROBINOW. The appearance of the electrocardiogram in relation to the position of the heart within the chest. Am. J. Med. Sci. **192:** 556, 1936.
(173) KATZ, L. N., E. SIGMAN, I. GUTMAN AND F. H. OCKO. The effect of good electrical conductors introduced near the heart on the electrocardiogram. Am. J. Physiol. **116:** 343, 1936.
(174) KATZ, L. N. AND A. W. WALLACE. The rôle of cardiac ischemia in producing R-T deviations in the electrocardiogram. Am. J. Med. Sci. **181:** 836, 1931.
(175) KATZ, L. N. AND S. F. WEINMAN. The relation of the T wave to the asynchronism between the ends of right and left ventricular ejection. Am. J. Physiol. **81:** 360, 1927.
(176) KAUFMAN, W. AND F. D. JOHNSTON. The electrical conductivity of the tissues near the heart and its bearing on the distribution of the cardiac action currents. Am. Heart J. **26:** 42, 1943.
(177) KELLOGG, O. P. Foundations of potential theory. F. Ungar Publ. Co., New York, 1943.
(178) KEITH, A. The sino-auricular node; a historical note. Brit. Heart J. **4:** 77, 1942.
(179) KIENLE, F. Untersuchungen über ein neues Verfahren zur Ableitung des Elektrocardiogramms von der Brustwand unter Anwendung einer differenten und einer indifferenten Elektrode. Arch. f. Kreislauff. **2:** 224, 1938.
(181) KISCH, B. The influence of the surface of heart muscle on the electrocardiogram. Cardiologia **4:** 304, 1940.
(182) KISCH, B. The changes of the chest leads of the electrocardiogram due to damage of the heart's surface. Cardiologia **4:** 318, 1940.
(183) KISCH, B., L. H. NAHUM AND H. E. HOFF. The predominance of surface over deep cardiac injury in producing changes in the electrocardiogram. Am. Heart J. **20:** 174, 1940.
(184) KISSIN, M., W. ACKERMAN AND L. N. KATZ. The effect of the heart's position on the electrocardiographic appearance of bundle branch block in man. Am. J. Med. Sci. **186:** 721, 1933.
(185) KOCH-MOMM, E. Experimenteller Beitrag zur Theorie des Elektrokardiogramms. Ztschr. f. Kreislauff. **24:** 501, 1932.
(186) KOCH-MOMM, E. Die Ungültigkeit des Einthoven'schen Dreieckschemas für das Elektrocardiogramm. Ztschr. f. Kreislauff. **25:** 513, 1933.
(187) KOCH-MOMM, E. Modellversuche über den Einfluss des elektrischen Feldes auf die Ableitung von Potentialdifferenzen. Ztschr. f. Biol. **93:** 241, 1933.
(188) KOREY, H. AND L. N. KATZ. The electrocardiographic changes produced by injuries of various parts of the ventricles. Am. J. Med. Sci. **188:** 387, 1934.

(189) KOSSMANN, C. E. AND F. D. JOHNSTON. The precordial electrocardiogram. I. The potential variations of the precordium and the extremities in normal subjects. Am. Heart J. **10**: 925, 1935.

(190) KOUNTZ, W. B. AND M. PRINZMETAL. Revival of the human heart. J. Clin. Investigation **13**: 711, 1934.

(191) KOUNTZ, W. B. AND M. PRINZMETAL. Influence of inversion of the dog on the electrocardiogram. Proc. Soc. Exper. Biol. and Med. **31**: 612, 1934.

(192) KOUNTZ, W. B., M. PRINZMETAL, E. F. PEARSON AND K. F. KOENING. The effect of position of the heart on the electrocardiogram. I. The electrocardiogram in revived perfused human hearts in normal position. Am. Heart J. **10**: 605, 1935.

(193) KOUNTZ, W. B., M. PRINZMETAL AND J. R. SMITH. The effect of position of the heart on the electrocardiogram. II. Observations upon the electrocardiogram obtained from a dog's heart placed in the human pericardial cavity. Am. Heart J. **10**: 614, 1935.

(194) KOUNTZ, W. B., M. PRINZMETAL AND J. R. SMITH. The effect of position of the heart on the electrocardiogram. III. Observations upon the electrocardiogram in the monkey. Am. Heart J. **10**: 623, 1935.

(195) KRASNO, M. R., J. A. E. EYSTER AND J. P. HETTWER. Action potentials from the heart of the early chick embryo. Am. J. Physiol. **119**: 357, 1937.

(196) KRASNO, M. R., J. A. E. EYSTER AND C. A. MAASKE. The nature of the T wave potentials in the tortoise heart. Am. J. Physiol. **114**: 119, 1935.

(197) LEPESCHKIN, E. Über das normale Brustwand-Elektrokardiogramm im Kindesalter. Arch. f. Kreislauff. **3**: 321, 1938.

(198) LEWIS, T. The mechanism and graphic registration of the heart beat. Shaw & Sons, Ltd., London, 1925, 3rd ed.

(199) LEWIS, T. AND M. A. ROTHSCHILD. The excitatory process in the dog's heart. II. Phil. Tr. London, Series B. **206**: 181, 1915.

(200) LINDNER, E. AND L. N. KATZ. The relative conductivity of the tissues in contact with the heart. Observations on animals with closed chests. Am. J. Physiol. **125**: 625, 1939.

(201) LLOYD, W. The form and function of the auriculo-ventricular bundle in the rabbit. Am. J. Anat. **45**: 379, 1930.

(202) LOMBARDINI, R. V. Une autre théorie sur l'onde T. Arch. Mal. du Coeur **26**: 548, 1933.

(203) LUNDY, C. J. AND C. M. BACON. Premature left ventricular beats from electrical stimulation of exposed human heart. Arch. Int. Med. **52**: 30, 1933.

(204) LUNDY, C. J., I. TREIGER AND R. DAVISON. Ventricular extrasystoles induced by electrical stimulation of the exposed human heart rotated 30° counterclockwise on its vertical axis. Am. Heart J. **17**: 85, 1939.

(205) MACLEOD, A. G. The electrogram of cardiac muscle. I. An analysis which explains the regression of T deflection. Am. Heart J. **15**: 165, 1938.

(206) MACLEOD, A. G. The electrogram of cardiac muscle. II. The lengths of the stages of activity. Am. Heart J. **15**: 402, 1938.

(207) MACLEOD, A. G., F. N. WILSON AND P. S. BARKER. The form of the electrocardiogram. I. Intrinsicoid electrocardiographic deflections in animals and man. Proc. Soc. Exper. Biol. and Med. **27**: 586, 1930.

(208) MAHAIM, I. Les maladies organiques du faisceau de His-Tawara. Masson & Cie., Paris, 1931.

(209) MAHAIM, I. Le syndrome de Wolff-Parkinson-White et sa pathogenie. Helvetica Med. Acta. **8**: 483, 1941.

(210. MAHAIM, I. AND M. R. WINSTON. Researches on the comparative anatomy and the experimental pathology of the superior connection of the bundle of His-Tawara. Cardiologia **5**: 189, 1941.

(211) MAHAIM, I., M. R. WINSTON AND H. ROESLER. Bilateral missed block, report of a case in which there was left bundle branch block during life with autopsy. Am. Heart J. **25**: 251, 1943.

(212) MANN, H. A method of analyzing the electrocardiogram. Arch. Int. Med. **25**: 283, 1920.

(213) MANN, H. The monocardiograph. Am. Heart J. **15**: 681, 1938.

(214) MARVIN, H. M. AND A. W. OUGHTERSON. The form of premature beats resulting from direct stimulation of the human ventricle. Am. Heart J. **7**: 471, 1932.

(215) MEEK, W. J., J. A. E. EYSTER AND H. GOLDBERG. Certain characteristics of injury potential in cardiac muscle. Am. J. Physiol. **126**: 583, 1939.

(216) MEHRING, C. E. AND K. E. ROTHSCHUH. Weiterer Untersuchungen über das elektrische Aktionsphenomena des Herzens. Ztschr. f. Biol. **100**: 68, 1940.

(217) MILLER, J. R. AND R. F. DENT. A new hypothesis of the production of the T wave in the electrocardiogram based on electro-kinetic phenomena. J. Clin. Investigation **19**: 783, 1940.

(218) MOLZ, B. Modellversuche zur kenntnis des elektrischen Feldes in der Elektrokardiographie. Zur Frage der Nullpotential-Elektrode und des Einthovenschen Dreieckschemas. Pflüger's Arch. **237**: 251, 1936.

(219) MOLZ, B. Eine Methode der unipolaren Ableitung des Elektrokardiogramms. Ztschr. f. Kreislauff. **29**: 361, 1937.

(220) MOLZ, B. Über die unipolare Ableitung des Elektrokardiogramms. Pflüger's Arch. **242**: 416, 1939.

(221) MÖNCKEBERG, J. G. Erkrankungen des Myocards und des spezifischen Muskelsystems. Handb. der sp. Path. Anat. und Histologies. J. Springer. Berlin, 1924.

(222) MYERS, G. B. AND B. G. OREN. The use of the augmented unipolar left leg lead in the differentiation of the normal from abnormal Q wave in standard lead III. Am. Heart J. **29**: 708, 1945.

(223) NACHMANSOHN, D. On the mechanism of transmission of nerve impulses. Collect. Net. **17**: August 1942.

(224) NACHMANSOHN, D., R. T. COX, C. W. COATES AND A. L. MACHADO. Phosphocreatine as energy source of the action potential. Proc. Soc. Exper. Biol. and Med. **52**: 97, 1943.

(225) NAHUM, L. H., W. F. HAMILTON AND H. E. HOFF. The injury current in the electrocardiogram. Am. J. Physiol. **139**: 202, 1943.

(226) NAHUM, L. H. AND H. E. HOFF. The recovery of electrical excitability in the mammalian ventricle. Am. J. Physiol. **126**: 589, 1939.

(227) NAHUM, L. H. AND H. E. HOFF. The interpretation of the U wave of the electrocardiogram. Am. Heart J. **17**: 585, 1939.

(228) NAHUM, L. H. AND H. E. HOFF. The localization of ventricular extrasystoles. Yale J. Biol. and Med. **17**: 539, 1945.

(229) NAHUM, L. H., H. E. HOFF AND W. KAUFMAN. Formation of the R complex of the electrocardiogram. Am. J. Physiol. **133**: 396, 1941.

(230) NAHUM, L. H., H. E. HOFF AND W. KAUFMAN. Configuration of anterior and posterior septal extrasystoles in the standard leads of the electrocardiogram. Am. J. Physiol. **134**: 398, 1941.

(231) NAHUM, L. H., H. E. HOFF AND W. KAUFMAN. Influence of temperature on the electrogram and monophasic action potential of the mammalian heart. Proc. Soc. Exper. Biol. and Med. **46**: 395, 1941.

(232) NAHUM, L. H., H. E. HOFF AND W. KAUFMAN. Formation of the R wave of the electrocardiogram. Am. J. Physiol. **134**: 384, 1941.

(233) NAHUM, L. H., H. E. HOFF AND W. KAUFMAN. The nature of the S complex of the electrocardiogram. Am. J. Physiol. **136**: 726, 1942.

(234) NAHUM, L. H., H. E. HOFF AND B. KISCH. The significance of displacement of the RS-T segment. Am. J. Physiol. **131**: 693, 1941.

(235) NONIDEZ, J. F. The structure and innervation of the conductive system of the heart of the dog and rhesus monkey, as seen with a silver impregnation technique. Am. Heart J. **26**: 577, 1943.

(236) NYBOER, J. Further observations on the esophageal electrocardiogram. J. Clin. Investigation **21**: 649, 1942.

(237) NYLIN, G. AND C. CRAFOORD. Simultaneous electrograms from left and right ventricles of the human heart. Cardiologia **6**: 136, 1942.

(238) NYLIN, G. AND K. E. GREWIN. The chest leads and their value in association with the leads from the extremities, and several other chest leads. Cardiologia **6**: 169, 1942.

(239) NYLIN, G. AND T. SÄLLSTRÖM. Three synchronized leads between fixed points on the heart projection on the chest wall. Acta med. Scand. **96**: 1, 1938.

(240) ÖHNELL, R. F. Pre-excitation, a cardiac abnormality. Acta med. Scand. Suppl. **152**: 1, 1944.

(241) PANNIER, R. Le tissue Purkinien des oreillettes du chat. Arch. Biol. **50**: 271, 1939.

(242) PANNIER, R. AND N. GOORMAGHTIGH. Sur l'existence de faisceaux Purkiniens dans l'oreillette droite. C. R. Soc. Biol. **127**: 1114, 1938.

(243) PAPP, C. U, the 6th wave of the electrocardiogram. Brit. Heart J. **2**: 9, 1940.

(244) PATTEN, B. M. Microcinematographic and electrocardiographic studies of the first heart beats and the beginning of the circulation in living embryos. Proc. Inst. Med. Chicago, **12**: 366, 1939.

(245) PATTEN, B. M. AND T. C. KRAMER. The initiation of contraction in the embryonic chick heart. Am. J. Anat. **53**: 349, 1933.

(246) PARDEE, H. E. B. AND M. GOLDENBERG. Electrocardiographic features of myocardial infarction as affected by involvement of the septum and by complete and incomplete transmural involvement. Am. Heart J. **30**: 367, 1945.

(247) PEARSON, H. E. S. AND L. M. GERLIS. A quantitative electrocardiographic method. Brit. Heart J. **7**: 173, 1945.

(248) POL, VAN DER, B. AND J. VAN DER MARK. Le battement du coeur considéré comme oscillation de relaxation et un modèle életrique due coeur. L'onde electrique **7**: 365, 1928 (also Arch. Neer. Physiol. **14**: 418, 1929).

(248a) PRINZMETAL, M., W. B. KOUNTZ AND D. P. BARR. The contour of the chest as a factor in the explanation of the difference between the dog and human electrocardiogram. J. Clin. Investigation **13**: 695, 1934.

(249) PRINZMETAL, M., W. B. KOUNTZ AND L. GOTTLIEB. Influence of point at which ventricle contacts thoracic wall on form of the electrocardiogram. Proc. Soc. Exper. Biol. and Med. **32**: 1411, 1935.

(250) PRINZMETAL, M., B. S. OPPENHEIMER AND S. DACK. Localization of ventricular extrasystoles. J.A.M.A. **108**: 620, 1937.

(251) PRUITT, R. D., A. R. BARNES AND H. E. ESSEX. Electrocardiographic changes associated with lesions in the deeper layers of the myocardium. Am. J. Med. Sci. **210**: 100, 1945.

(252) PUDDU, V. AND C. CAMMARELLA. Un cas d'ectopia cordis avec une étude electrocardiographique de la progression de l'onde d'excitation. Arch. Mal. du Coeur **31**: 861, 1938.

(253) RAPPORT, D. AND G. B. RAY. Changes of electrical conductivity in the beating tortoise ventricle. Am. J. Physiol. **80**: 126, 1927.

(254) RAWLINSON, H. E. The later development of the sinus venosus and the relation of the sino-atrial node to it, in the calf heart. Anat. Rec. **49**: 89, 1931.

(255) RIJLANT, P. Contribution à l'étude de l'automatisme et de la conduction dans le coeur. II. Conduction intra-auriculaire chez le mammifère. Arch. Intern. Physiol. **27**: 225, 1926.

(256) RIJLANT, P. La conduction dans l'oreillette droite du coeur du mammifère. C. R. Soc. Biol. **103**: 909, 1930.

(257) RIJLANT, P. La conduction dans le coeur du mammifère. Ann. Physiol. et Physico-Chim. Biol. **7**: 229, 1931.

(258) RIJLANT, P. La conduction dans le coeur du mammifère. Arch. Internat d. Physiol. **33**: 325, 1931.

(259) Rijlant, P. Mécanisme de l'envahissement de l'oreillete droite du coeur du mammifère par la contraction. C.R. Soc. Biol. **121**: 1361, 1936.
(260) Rijlant, P. Introduction à l'étude de la distribution spatiale des variations de potentiel produites par le coeur chez l'homme. C.R. Soc. Biol. **121**: 1358, 1936.
(261) Robb, J. S. Evidence for characteristic modifications of the electrocardiogram produced by lesions of ventricular muscle bands. Proc. Soc. Exper. Biol. and Med. **31**: 311, 1933.
(262) Robb, J. S. The structure of the mammalian auricle. Med. Women's J. **41**: 65, 1934.
(263) Robb, J. S. The distribution of the Purkinje material in the septum of the beef heart. Am. J. Physiol. **126**: 608, 1939.
(264) Robb. J. S. A study of Q-T interval in various species. Fed. Proc. **5**: 87, 1946.
(265) Robb, J. S., M. S. Dooley and R. C. Robb. Displacement of the RS-T segment by potassium chloride. J. Mt. Sinai Hosp. **8**: 946, 1942.
(266) Robb, J. S., M. Easby and J. G. F. Hiss. Experimental interference with conduction in the heart. Am. J. Med. Sci. **188**: 835, 1934.
(267) Robb, J. S., W. Greene and R. C. Robb. The peripheral distribution of the Purkinje fibres. J. Tech. Methods **17**: 91, 1937.
(268) Robb, J. S. and J. G. F. Hiss. Experimental interference with conduction in the ventricle Am. J. Physiol. **109**: 89, 1934.
(269) Robb, J. S., J. G. F. Hiss and R. C. Robb. Localization of cardiac infarcts according to component ventricular muscles. Am. Heart J. **10**: 287, 1935.
(270) Robb, J. S. and C. T. Kaylor. The A-V conduction system in the heart of the guinea pig. Proc. Soc. Exper. Biol. and Med. **59**: 92, 1945.
(271) Robb, J. S. and R. C. Robb. The excitatory process in the mammalian ventricle. Am. J. Physiol. **115**: 43, 1936.
(272) Robb, J. S. and R. C. Robb. The conducting system. Am. J. Physiol. **119**: 390, 1937.
(273) Robb, J. S. and R. C. Robb. Converse effects of stimulating opposite ends of a cardiac muscle band. Proc. Soc. Exper. Biol. and Med. **37**: 187, 1937.
(274) Robb, J. S. and R. C. Robb. The S wave of the electrocardiogram. Am J. Physiol. **126**: 608, 1939.
(275) Robb, J. S. and R. C. Robb. Hypertension electrocardiograms experimentally produced and anatomically explained. I. Cor pulmonale. Am. J. Med. Sci. **203**: 625, 1942.
(276) Robb, J. S. and R. C. Robb. Hypertension electrocardiograms experimentally produced and anatomically explained. II. Left ventricular strain. Am. J. Med. Sci. **203**: 634, 1942.
(277) Robb, J. S., R. C. Robb and J. G. F. Hiss. Localization of premature beats in the mammalian ventricle. Proc. Soc. Exper. Biol. and Med. **32**: 1510, 1935.
(278) Robb, R. C. and J. S. Robb. Experimental bundle branch block after ablation of one or both ventricles. Am. J. Med. Sci. **204**: 313, 1942.
(279) Robb, J. S. and R. C. Robb. The normal heart. Anatomy and physiology of the structural unit. Am. Heart J. **23**: 455, 1942.
(280) Robberts, G. H., J. C. Hamilton and D. I. Abramson. Experimental bundle branch block in the monkey. J. Clin. Investigation **14**: 867, 1935.
(281) Robinow, M. A proved case of dextroposition of the heart showing left axis deviation in the electrocardiogram. Am. Heart J. **13**: 104, 1937.
(282) Robinow, M., L. N. Katz and A. Bohning. The appearance of the T wave in lead IV in normal children and in children with rheumatic heart disease. Am. Heart J. **12**: 88, 1936.
(283) Rochet, J. and M. M. Vastesaeger. *Le vectocardiogramme de l'homme normal,* Travaux Labor. Inst. Solvay Physiol. **29**: 17, 1944.
(284) Rojas, P. and J. Debussy-Schleisinger. El tejido conectivo de los distintos segmentos del sistema de conducción aurículo-ventricular. Rev. Soc. Argent. Biol. **12**: 313, 1936.

(285) Rosenbaum, F. F., F. N. Wilson and F. D. Johnston. The precordial electrocardiogram in high lateral myocardial infarction. Proc. Cent. Soc. Clin. Res. 18: 35, 1945.

(286) Rosenblueth, A., W. Daughaday and D. D. Bond. The action of electrial stimuli on the turtle's ventricle. Am. J. Physiol. 138: 50, 1942.

(287) Rosenblueth, A., W. Daughaday and D. D. Bond. The electrogram of the ventricle of the turtle's heart. Am. J. Physiol. 139: 464, 1943.

(288) Rosenblueth, A. and E. C. del Pozo. The changes of impedance of the turtle's ventricular muscle during contraction. Am. J. Physiol. 139: 514, 1943.

(289) Rosenblueth, A. and J. Garcia-Ramos. La interpretación del electrograma del ventrículo cardiaco. Arch. Inst. Cardiologica México 14: 232, 1945.

(290) Rosenblueth, A. and J. Garcia-Ramos. La acción de la corriente directa sobre los diversos componentes del electrograma del ventrículo cardíaco. Arch. Inst. Cardiologica México 15: 1, 1945.

(291) Rosenblueth, A. and J. García-Ramos. La acción del calor y el frío sobre los diversos componentes del electrograma del ventrículo cardíaco. Arch. Inst. Cardiologica México 15: 101, 1945.

(292) Rosenblueth, A. and J. Garcia-Ramos. Los diversos componentes del electrograma monofásico del ventrículo cardíaco. Arch. Inst. Cardiologica México 16: 45, 1946.

(293) Rothberger, C. J. Beiträge zur normalen and pathologischen Physiologie der spezifischen Herzmuskulatur. Cardiologia 1: 234, 1937.

(294) Rothberger, C. J. Normale and pathologische Physiologie der Rhythmik und Koordination des Herzens. Ergbn. d. Physiol. 32: 472, 1931.

(295) Rothschuh, K. E. Über die Entstehung von Verletzungsströmen und monophasischer Derformierung der Herzaktions-Ströme durch Körpereigene Substanzen. Ztschr. f. exper. Med. 106: 543, 1939.

(296) Rothschuh, K. E. Über die Entstehung von Q- and S-Zacken im Elektrokardiogram. Pflüger's Arch. 246: 820, 1943.

(297) Routier, D. Le vectogramme. Arch. Mal. du Coeur 31: 697, 1938.

(298) Sanabria, T. Recherches sur la differenciation du tissu nodal et connecteur du coeur des Mammifères. Arch. Biol. 47: 1, 1935.

(299) Schaefer, H. Elektrophysiologie, II. Band: Spezielle Elektrophysiologie; F. Deuticke, 1942, Lithoprinted Edward Bros. Inc.; Ann Arbor, 1944, pp. 1 to 130 and pp. 331 to 348.

(300) Schaefer, H. Theorie des Potentialabgriffs beim Elektrokardiogramm auf der Grundlage der Membran-Theorie. Pflüger's Arch. 245: 72, 1941.

(301) Schaefer, H., A. Pena and P. Schölmerich. Der monophasiche Aktionsstrom der Spitze und Basis dos Warmblüterherzens und die Theorie der T Welle des Ekg. Pflüger's Arch. 246: 728, 1943.

(302) Schellong, F. Vektordiagraphie des Herzens als klinische Methode. Klin. Wschr. 17: 453, 1938.

(303) Schellong, F., S. Heller and E. Schwinge. Das Vektordiagramm; eine Untersuchungs-Methode des Herzens. Ztschr. f. Kreislauff. 29: 497, 1937.

(304) Schellong, F. and E. Schwinge. Das Vektordiagramm; eine Untersuchungsmethode des Herzens; über die Bedeutung von Knotungen und Aufsplitterung in QRS des Ekg. Ztschr. f. Kreislauff. 29: 596, 1937.

(305) Schellong, F., E. Schwinge and H. Hermann. Die praktisch-klinische Methode der Vektordiagraphie und das normale Vektordiagramm. Arch. f. Kreislauff. 2: 1, 1937.

(306) Schocken, K. The validity of Einthoven's triangle rule. Cardiologia 3: 197, 1939.

(307) Schütz, E. Monophasische Aktionsströme vom in situ durchbluteten Säugetierherzen. Klin. Wchnschr. 31: 1454, 1931.

(308) Schütz, E. Einphasische Aktionsströme vom in situ durchbluteten Säugetierherzen. Ztschr. f. Biol. 92: 441, 1932.

(309) Schütz, E. Elektrophysiologie des Herzens bei einphasischer Ableitung. Ergebn. d. Physiol. 38: 493, 1936.

(310) Schütz, E., K. E. Rothschuh and C. E. Mehring. Ein direkter Beweis für die Gültigkeit der Differenzkonstruktion des Elektrogramms. Klin. Wchnschr. 19: 9, 1940.

(311) Segers, M. Le parallélisme entre l'accommodation de l'excitabilité et du rythme du coeur. Acta Biol. Belg. 1: 255, 1941.

(312) Segers, M. L'identité de l'onde U et de la négativité tardive du coeur. Acta Biol. Belg. 3: 72, 1943.

(313) Shaner, R. F. The development of the atrio-ventricular node, bundle of His, and sino-atrial node in the calf, with a description of a third embryonic node-like structure. Anat. Rec. 44: 85, 1929.

(314) Sodi Pallares, D. Nuevas bases de electrocardiografia. Talleres Gráficos de la Nación, Mexico, D. F., 1945, Ediciones del Instituto Nat. de Cardiología.

(315) Sodi Pallares, D., A. Cuellar and E. Cabrera. Sistema de seis ejes con aplicación al vector AvT en las hipertrofias ventriculares. Arch. Inst. Cardiologica México 14: 142, 1945.

(316) Stratton, J. A. Electromagnetic theory. McGraw-Hill Book Co., N.Y.C., 1941.

(317) Storti, E. Zur Frage der Nullpotential Elektrode in der Elektrokardiographie. Ztschr. f. Kreislauff. 27: 830, 1935.

(318) Sugarman, H., L. N. Katz, A. Sanders and K. Jochim. Observations on the genesis of the electrical currents established by injury to the heart. Am. J. Physiol. 130: 130, 1940.

(319) Sulzer, R. and P. W. Duchosal. L'électrocardiographie à deux dimensions ou planélectrocardiographie. Helvet. Med. Acta. 4: 740, 1937.

(320) Sulzer, R. and P. W. Duchosal. Application de la planographie à l'étude de la distribution des potentiels du coeur à la surface du corps. Arch. Mal. d. Coeur 31: 686, 1938.

(321) Sulzer, R. and P. W. Duchosal. Die Grundform des menschlichen Vektorkardiogramms. Schweiz. Med. Wchnschr. 71: 259, 1941.

(322) Sulzer, R. and P. W. Duchosal. Principle of cardiovectography. Cardiologia 6: 236, 1942.

(323) Sulzer, R. and P. W. Duchosal. Le cardiovectogramme normal. Helv. Physiol. Acta. 4: 285, 1946.

(324) Szepesenwol, J. and J. B. Odoriz. Potenciales de acción de esbozo cardíaco de embriones de pollo y de rata cultivádos in vitro. Rev. Soc. Argent. Biol. 19: 279, 1943.

(325) Taussig, H. B. Boundaries of sino-auricular node and atrio-ventricular node in human heart. Johns Hopkins Hosp. Bull. 48: 162, 1931.

(326) Ter Borg, H. (Purkinje tissue in the atria of domestic animals (esp. the horse)). Acta. Neer. Morph. 1: 64, 1937.

(327) Tigerstedt, R. Die Physiologie des Kreislaufes. W. de Gruyter, Berlin, 1921. (4 vol.)

(328) Vastesaeger, M. M. and J. Rochet. Les propriétés du vectocardiogramme spatial de l'homme normal, ses variations physiologiques. Trav. Lab. Inst. Solvay Physiol. 29: 55, 1944.

(329) Vastesaeger, M. M. and J. Rochet. La stereovectocardiographie et la stereovectocardioscopie, méthodes cliniques d'étude de la repartition spatiale des potentiels cardiaques. Trav. Lab. Inst. Solvay Physiol. 29: 40, 1944.

(330) Wahlin, B. Das Reizleitungssystem und die Nerven des Säugetierherzens. Inaug. Dissert. Stockholm, 1935, also Sv. Lak. Sallsk. Hdl. 62: 1, 1936.

(331) Waller, A. On the electromotive changes connected with the beat of the mammalian heart, and of the human heart in particular. Phil. Trans. Royal Soc. London, Series B. 180: 169, 1889.

(332) Walls, E. W. Specialised conducting tissue in heart of golden hamster. J. Anat. 76: 359, 1942.

(333) Walls, E. W. Specialised conducting tissue in heart of common hedgehog. J. Anat. 77: 294, 1943.

(334) Walls, E. W. Dissection of atrio-ventricular node and bundle in human heart. J. Anat. 79: 45, 1945.

(335) Weber, A. Klinische und experimentelle Studien über das Elektrokardiogramm; Experimentelle Untersuchungen zur Deutung des Elektrokardiogramms. Ztschr. klin. Med. 132: 153, 1937.

(336) Wedd, A. M. and W. D. Stroud. The spread of the excitation wave related to the standard electrocardiogram in the dog's heart. Heart 9: 15, 1921.

(337) Weinstein, J. and D. I. Abramson. Localization of the site of experimental premature contractions and bundle branch lesions by means of multiplane chest leads. Ann. Int. Med. 9: 1187, 1936.

(338) Wenckebach, K. F. and H. Winterberg. Die unregelmässige Herztätigkeit. W. Engelmann, Leipzig, 1937.

(339) Wiggers, C. J. The independence of electrical and mechanical reactions in the mammalian heart. Am. Heart J. 1: 173, 1925.

(340) Wiggers, C. J. Monophasic and deformed ventricular complexes resulting from surface applications of potassium salts. Am. Heart J. 5: 346, 1930.

(341) Wiggers, C. J. and M. G. Banus. A. On the independence of electrical and mechanical activity in the mammalian ventricle. B. On the effects of pH changes on conduction in the heart. Am. J. Physiol. 76: 215, 1926.

(342) Wiggers, H. C. Pure monophasic action potentials and their employment in studies of ventricular surface negativity. Proc. Soc. Exper. Biol. and Med. 34: 337, 1936.

(343) Wiggers, H. C. The sequence of ventricular surface excitation determined by registration of monophasic action potentials. Am. J. Physiol. 118: 333, 1937.

(344) Wiggers, H. C. The time relationship of electrical and mechanical processes in the ventricle of the dog. Am. J. Physiol. 123: 215, 1938.

(345) Wiggers, H. C. and C. J. Wiggers. The interpretation of monophasic action potentials from the mammalian ventricle indicated by changes following coronary occlusion. Am. J. Physiol. 113: 683, 1935.

(346) Wilde, W. S., C. E. Drawe, and R. Ashman. The sequence of excitation of the surface of the turtle ventricle. Am. J. Physiol. 129: 497, 1940.

(347) Wilson, F. N. The distribution of potential differences produced by the heart beat within the body and at its surface. Am. Heart J. 5: 509, 1930.

(348) Wilson, F. N., P. S. Barker, F. D. Johnston, I. G. W. Hill and G. Grout. The electrocardiogram in the earlier stages of experimental myocardial infarction. J. Clin. Investigation 12: 993, 1933.

(349) Wilson, F. N., I. G. W. Hill and F. D. Johnston. The form of the electrocardiogram in experimental myocardial infarction. Am. Heart J. 9: 596, 1934.

(350) Wilson, F. N., I. G. W. Hill and F. D. Johnston. The form of the electrocardiogram in experimental myocardial infarction. III. Am. Heart J. 10: 903, 1935.

(351) Wilson, F. N. and F. D. Johnston. The vectorcardiogram. Am. Heart J. 16: 14, 1938.

(352) Wilson, F. N., F. D. Johnston, I. G. W. Hill and G. Grout. Form of the electrocardiogram in experimental myocardial infarction. Proc. Soc. Exper. Biol. and Med. 30: 799, 1933.

(353) Wilson, F. N., F. D. Johnston and I. G. W. Hill. The form of the electrocardiogram in experimental myocardial infarction. IV. Am. Heart J. 10: 1025, 1935.

(354) Wilson, F. N., F. D. Johnston and P. S. Barker. The use of the cathode-ray oscillograph in the study of the monocardiogram. J. Clin. Investigation 16: 664, 1937.

(355) Wilson, F. N., F. D. Johnston, A. G. Macleod and P. S. Barker. Electrocardiograms that represent the potential variations of a single electrode. Am. Heart J. 9: 447, 1934.

(356) Wilson, F. N., F. D. Johnston, F. F. Rosenbaum and P. S. Barker. On Ein-

thoven's triangle, the theory of unipolar electrocardiographic leads, and the interpretation of the precordial electrocardiogram. Am. Heart J. 32: 277, 1946.
(357) Wilson, F. N., F. D. Johnston, F. F. Rosenbaum, H. Erlanger, C. E. Kossmann, H. H. Hecht, N. Cotrim, R. M. de Oliveira, R. Scarsi and P. S. Barker. The precordial electrocardiogram. Am. Heart J. 27: 19, 1944.
(358) Wilson, F. N., A. G. Macleod and P. S. Barker. The form of the electrocardiogram. II. The character of the excitation wave in auricular muscle. Proc. Soc. Exper. Biol. and Med. 27: 588, 1930.
(359) Wilson, F. N., A. G. Macleod and P. S. Barker. The form of the electrocardiogram. III. Opposed potential difference. Proc. Soc. Exper. Biol. and Med. 27: 589, 1930.
(360) Wilson, F. N., A. G. Macleod and P. S. Barker. The form of the electrocardiogram. IV. The mean electrical axis and the center of stimulation. Proc. Soc. Exper. Biol. and Med. 27: 590, 1930.
(361) Wilson, F. N., A. G. Macleod and P. S. Barker. The interpretation of the initial deflections of the ventricular complex of the electrocardiogram. Am. Heart J. 6: 637, 1931.
(362) Wilson, F. N., A. G. Macleod and P. S. Barker. The accuracy of Einthoven's equation and the potential variations produced by the heart beat at the apices of Einthoven's triangle. Am. Heart J. 7: 203, 1931.
(363) Wilson, F. N., A. G. Macleod and P. S. Barker. Potential variations produced by the heart beat at the apices of Einthoven's triangle. Am. Heart J. 7: 207, 1931.
(364) Wilson, F. N., A. G. Macleod and P. S. Barker. The order of ventricular excitation in human bundle-branch block. Am. Heart J. 7: 305, 1932.
(365) Wilson, F. N., A. G. Macleod and P. S. Barker. Electrocardiographic leads which record potential variations produced by the heart beat at a single point. Proc. Soc. Exper. Biol. and Med. 29: 1011, 1932.
(366) Wilson, F. N., A. G. Macleod and P. S. Barker. The distribution of the action currents produced by heart muscle and other excitable tissues immersed in extensive conducting media. J. Gen. Physiol. 16: 423, 1933.
(367) Wilson, F. N., A. G. Macleod and P. S. Barker. The distribution of the currents of action and of injury displayed by heart muscle and other excitable tissue. Univ. Michigan Press, Ann Arbor, 1933.
(368) Wilson, F. N., A. G. Macleod, P. S. Barker and F. D. Johnston. The determination and the significance of the areas of the ventricular deflections of the electrocardiogram. Am Heart J. 10: 46, 1934.
(369) Wilson, F. N., A. G. Macleod, F. D. Johnston and I. G. W. Hill. Monophasic electrical response produced by the contraction of injured heart muscle. Proc. Soc. Exper. Biol. and Med. 30: 797, 1933.
(370) Wilson, F. N., S. W. Wishart and G. R. Herrmann. Factors influencing distribution of potential differences, produced by heart-beat, at surface of body. Proc. Soc. Exper. Biol. and Med. 23: 276, 1926.
(371) Wolferth, C. C. and M. M. Livezey. A study of methods of making so-called unipolar electrocardiograms. Am. Heart J. 27: 764, 1944.
(372) Wolferth, C. C., M. M. Livezey and F. C. Wood. The relationships of lead I, chest leads from the C_3, C_4 and C_6 positions, and certain leads made from each shoulder region: the bearing of these observations upon the Einthoven equilateral triangle hypothesis and upon the formation of lead I. Am. Heart J. 21: 215, 1941.
(373) Wolferth, C. C., M. M. Livezey and F. C. Wood. Studies on the distribution of a potential of ventricular origin. Am. J. Med. Sci. 205: 154, 1943.
(374) Wolferth, C. C., M. M. Livezey and F. C. Wood. Distribution of the patterns of ventricular potential which determine the forms and significance of electrocardiograms. Am. J. Med. Sci. 205: 469, 1943.

(375) Wolferth, C. C., S. Bellet, M. M. Livezey and F. D. Murphy. Negative displacement of the RS-T segment in the electrocardiogram and its relationship to positive displacement. Am. Heart J. **29**: 220, 1945.
(376) Wolferth, C. C. and F. C. Wood. The prediction of differences between precordial leads CR, CL, and CF based on limb lead findings. Am. Heart J. **20**: 12, 1940.
(377) Wood, F. C., C. C. Wolferth and G. D. Geckeler. Histologic demonstration of accessory muscular connections between auricle and ventricle in a case of short P-R interval and prolonged QRS complex. Am. Heart J. **25**: 454, 1943.
(378) Yater, W. M., V. H. Cornell and T. Claytor. Auriculoventricular block due to bilateral bundle branch lesions. Arch. Int. Med. **57**: 132, 1936.
(379) Yater, W. M. Pathogenesis of bundle branch block. Arch. Int. Med. **62**: 1, 1938.
(380) v. Zarday, I. Ekg-Axonometer. Ein neuer Apparat zur Bestimmung der elektrischen Achsen des Herzens. Klin. Wchnschr. **19**: 1084, 1940.
(381) v. Zarday, I. Bedeutung der elektrischen Momentanachsen des Herzens. Arch. f. Kreislauff. **7**: 223, 1940.
(382) v. Zarday, I. Das Elektro-Kardio-Axonogramm. Ztschr. klin. Med. **140**: 514, 1942.
(383) v. Zarday, I. Ein Beitrag zur Lösung des Vektordiagramms. Ztschr. f. Kreislauff. **35**: 18, 1943.
(384) Zeisler, E. B. A critique of Einthoven's law in electrocardiography. Proc. Soc. Exper. Biol. and Med. **28**: 12, 1930.
(385) Zeisler, E. B. and L. N. Katz. Studies of the electrical field of the heart. I. Invariants of the electrocardiogram. Am. Heart J. **8**: 676, 1933.

INTERCELLULAR CEMENT AND CAPILLARY PERMEABILITY

ROBERT CHAMBERS AND B. W. ZWEIFACH[1]

Department of Biology, Washington Square College of Arts and Sciences, New York University, New York

The passage of materials across a membrane consisting of cells, e.g., the capillary wall, requires a consideration of several structural and environmental factors. Environmental effects may be mediated through chemical action on the structure of the membrane; or they may be mechanical through the exertion of pressure on the membrane. The effects of the mechanical factor may be purely a matter of degree, variations in the amount of material passed through in a given time occurring without changing the permeability characteristics of the membrane; or they may be qualitative, distending and thereby increasing the porosity of the membrane. Even when considering the interrelations of purely extrinsic factors such as hydrostatic and osmotic pressure, it is necessary to take into account both qualitative and quantitative effects on the fluid-exchange across the membrane. The action of specific chemical substances obviously involves qualitative changes. In this regard an understanding of the structural components of the membrane is of prime importance. Moreover, since we are dealing with a hemodynamic system, we must treat the permeability of the membrane in relation to the multiplicity of factors which control the peripheral distribution of the blood.

On the basis of structural factors conditioning the passage of material across a cellular membrane we may distinguish three types (1). The permeability of one depends entirely on the cells constituting the membrane. This type is exemplified by secretory epithelium in which the constituent cells perform work in selectively absorbing material from the environment on the one side, passing it through the cytoplasm, and expelling it to the environment on the other side of the cell. An example of this kind is the wall of the renal proximal tubule in which a one-way permeability is so pronounced that even water is driven through in one direction against an osmotic pressure gradient (2). The hydrostatic pressure developed within such a tubule may be so high that the tubule distends and the constituent cells become compressed from their normal cuboidal to a flat pavement shape. Moreover, this can occur without affecting the one-way, selective permeability of the membrane.

Any cellular membrane which depends for its permeability on its constituent cells must be highly selective. The life of a cell is conditioned by its ability to maintain a differential between the environment of the cell and its interior. Concerning the penetration of substances into and their escape out of an individual cell, much investigation has been done and speculation advanced. But even the speculations have never gone so far as to assume that a cell can serve as

[1] Present address: Dept. of Medicine, Cornell University Medical College, New York, New York.

an indiscriminate thoroughfare for all material existing within certain critical dimensions of size. If the passage of materials into a cell, through its cytoplasm and out to the other side, is a sole property of colloidal and hydrostatic pressure differences, then our whole concept concerning cell-life falls to the ground.

The second type of cellular membrane has a permeability in which both the constituent cells and the intercellular material take part. Hoeber (3) has presented evidence for such an occurrence in the intestinal mucosa. In this type, the secretory activity of the cellular constituents predominates so that the composition of the fluid passed through the membrane is a mixture of secretion and intercellular diffusion products.

The third type of cellular membrane is one in which the permeability depends primarily on the intercellular cement-substance. The rôle of the cells would then be that of building blocks and be responsible mainly for elaborating the cement which lies between them and holds them together.

The relatively undiscriminating nature of the permeability of the wall of the blood capillary classifies the endothelium as a cellular membrane of this third type.[3] In other words, the permeability rôle of the constituent cells would be decidedly subordinate to that of the intercellular substance, otherwise the endothelial cell would have to be so porous as to make it difficult to conceive how such a body of protoplasm can maintain the highly specific attributes of a living cell (cf. Krogh (4)).

The mere fact of the thinness of the endothelial cell does not necessarily indicate that it possesses the wide range of permeability of the capillary wall. Layers of living protoplasm as thin as the endothelium possess a high degree of selectivity in their permeability to substances in solution. For example, the protoplast surrounding the vacuole of a marine plant cell, Valonia, is only a few microns thick, nevertheless it maintains a pronounced degree of ionic difference between the fluid in the vacuole and the surrounding medium.

For water and for nonelectrolytes of sufficiently small dimensions, e.g., CO_2 and O_2, passage into and through a cell may occur by physical diffusion depending on concentration gradients and differences of osmotic and hydrostatic forces on the two opposite sides of the cell. On the other hand, the cellular transport of electrolytes has been recognized only in secreting cells through a mechanism involving expenditure of energy beyond the basal requirements for vital integrity. The endothelial cell may serve in the physical transfer of water and presumably of dissolved gases but the permeability of the endothelial wall for substances such as all electrolytes, large nonelectrolytes and colloidal suspensions is, except

[3] There are special instances when the composition of tissue or body fluids in a particular regional area cannot be explained by simple ultra-filtration through the wall of blood vessels, e.g., synovial fluid, aqueous humor, cerebro-spinal fluid, extracellular fluid of brain tissue, etc. In these cases the barrier between the blood and the particular body cavity is double, consisting of an endothelial covered by an epithelial layer. Since the evidence overwhelmingly indicates that the endothelial layer acts as a simple ultrafilter, a secretory mechanism of the epithelial layer is probably responsible for the selective permeability of the barrier.

for particle size, so non-selective that the cell physiologist would prefer to relegate their passage to some constituent of the wall other than the living endothelial cells.

The alternative is the unlikely occurrence of active uptake by cells through a complex process of phagocytosis on an ultra-microscopic scale. The uptake by living cells of particulate matter is well recognized as phagocytosis. This process of ingestion has been extended to the uptake of colloidal dyes by means of ultramicroscopic phagocytosis. On the other hand, the only known mechanism for the elimination of such particles is either by a dissolution of the cell containing the particles or by a pinching-off of a bleb of cytoplasm enclosing one or more particles. This must involve expenditure of energy and a complex mechanism. It would be highly uneconomical for endothelial cells to use this method routinely in transferring particles from the lumen of the blood vessel to the outside. Florey (5) injected starch grains into the blood stream and found some lying in the endothelium and others outside. His claim for their passage through the endothelial cell is by inference only. Field and Drinker (6) observed the passage of calcite particles through the endothelial wall. The slowness of the passage led them to postulate uptake by the endothelial cell from the blood and ejection on the other side of the wall. However, passage through spaces in a cement substance, which presumably is sticky, may also be slow. Moreover, the physical consistency of the cement substance between the cells must be such as to permit the passage of particles of colloidal size up to aggregates and even leucocytes and red blood cells.

The material between the cells can act as a physical filter without the expenditure of energy, a mode of filtration stressed by Landis (11). Such a filter could exhibit variations in its limiting porosity of vessels in different tissues and also in the same tissue under different environmental conditions. Likewise, the findings of Abell (7) that newly developing blood capillaries are more highly permeable than later, suggest the gradual development of an interendothelial cement. Wilbrandt (8) cites the limiting porosity of the capillaries in the glomerular tuft of the kidney to be that of a molecular volume of 68,000. The uppermost limit of porosity of a cell membrane has never been claimed to be in excess of 342, the molecular volume of saccharose. A route other than the cell proper should be postulated in order to account for the passage, in bulk, of molecular aggregates through the endothelial wall.

The term, capillary permeability, has been loosely used without discriminating between those phenomena which are indirect in their action, i.e., through forces in the blood and surrounding tissues which do not affect the intrinsic property of the membrane barrier itself, and those which are direct, i.e., through factors which induce actual changes in the composition of the capillary wall.

A search of the literature reveals the lack of clear-cut evidence for defining the type of permeability referable to the changes being discussed. Many of the fluctuations in tissue-blood exchange which have been described as evidence for altered capillary permeability may not involve actual changes in the structure of the capillary wall. Such fluctuations are to be attributed rather to extrinsic

phenomena, such as fluctuations in hydrostatic pressure, factors regulating peripheral blood flow, avidity of extra-capillary tissue for either solute or solvent, under normal or experimentally induced conditions, etc. Alterations in permeability to be ascribed to changes in the membrane would be of primary importance in pathological phenomena. Under normal conditions it is highly probable that factors exist which maintain the capillary wall in a stable physico-chemical state. An elucidation of these factors must distinguish between effects on different components of the capillary wall.

ARCHITECTURE OF THE CAPILLARY WALL. There is good evidence for the existence of three structural components of the capillary wall, variations of each of which may alter conditions affecting the passage of materials through the wall. These are: the endothelium *per se*, consisting of pavement-like cells, and an intercellular cement; an endocapillary lining, which is non-cellular and is possibly derived from the circulating blood proteins; and third, a pericapillary sheath, which serves as an outer, supporting layer with characteristics common to the surrounding connective tissue matrix.

1. *The Capillary Endothelium.* The outstanding component of the capillary wall is the endothelium proper, the barrier usually considered as regulating the passage of materials from blood to the surrounding tissue. Most discussions assume that the exchange across the endothelial membrane involves the endothelial cells and Danielli (9) has gone so far as to postulate a continuous plasma membrane as the conditioning factor in capillary permeability. A more likely alternative, based on the highly porous character of the capillary wall, is to relegate a prominent rôle to the intercellular cement acting as physical ultrafilter.

Insistence on the cell as the primary route for tissue-blood fluid exchange would necessitate considering the endothelial cell as a radically modified, parchment-like cell remnant, an unwarranted assumption since the endothelial cell is known to possess fundamental properties of a living cell. It is capable of changing its form and size; it possesses tone; it is irritable and reacts to prodding; it undergoes mitotic division and, in common with living cells, its nucleus is a prominent structure which can be stained with basic dyes only after death. The concept that the endothelial cell plays a subordinate rôle in the exchange across the capillary wall does not imply that the cells are not vitally concerned in maintaining the capillary wall as a living structure. Aside from their functions in repair and regeneration there is evidence (10) that the endothelial cell elaborates the cement substance which binds the cells into a continuous membrane.

The concept of capillary permeability presented in this review is based on the properties of the non-cellular components of the capillary wall. The intercellular cement substance serves as a filter, the selective properties of which depend upon variations in its porosity. The passage of material through such a membrane would depend upon the number, size, shape, distribution and other properties of the pores of the cement also upcn the nature of their coating and the extent to which the pores can be plugged.

A quantitative consideration of the passage of substances through the capil-

lary wall must take into account the surface area of the cement relative to that of the cells available for filtration. Landis (11) cites figures to show that the filtration rate across the capillary wall occurs at 3.7×10^{-3} cc./cm.^{2}/min./atmos. This is approximately 100 times the value for the permeability to water of the arbacia egg as computed by Lucké. Wilbrandt (cf. 8) made a comparison of the absolute water permeability of red blood cells, leucocytes, fibroblasts, arbacia eggs and amoebae with that of a blood capillary wall. On the basis of equal pressure, area and time, he arrived at the same value as that of Landis, i.e., the permeability of the blood capillary to be 100$\times$ that of a living cell. Let us assume that the endothelial wall is 99 per cent cellular and 1.0 per cent intercellular and let us also assume that the permeability of the individual endothelial cell is the same as that of cells which have been measured. Then in 1 sq. cm. of capillary wall we would have 0.01 sq. cm. of intercellular cement having a high permeability and 0.99 sq. cm. surface of the cellular constituent having a low permeability. The permeability constant as found by Landis could refer to only the one hundredth part of the square centimeter wall consisting of cement substance. Let us imagine 1 sq. cm. of wall consisting entirely of the cement substance. The permeability constant of this 1 sq. cm. area would be larger by a value of 4 orders of magnitude than a comparable area consisting only of cells. Thus, any slight change in the nature of the cement should enormously affect the permeability of the endothelial wall. The pore size in the cement may vary under different conditions, e.g., mechanical stretching of the membrane or by changes in the chemical nature and pH of surrounding medium. Such changes may influence the amount of fluid which passes in a given time. Furthermore, a mere change in shape of the pores from an elliptical to a circular, will change the limiting porosity of the membrane to colloidal constituents without affecting the rate of water exchange (12).

Intercellular cement. In view of the relative importance of the inter-endothelial cement it is necessary to discuss factors concerned with the maintenance of its physical state. The early investigators, Ringer, Herbst, Overton, stressed the existence of a reversible, organic calcium salt which serves as a cohesive substance for binding cells together. The physical state of the cement can be varied by changing either the pH or the calcium content of the medium (cf. 10). This non-cellular constituent of the capillary wall was recognized by Cohnheim in his earlier papers about 1867 and by Arnold in 1887 who claimed that a loosening of the cement substance may become so great as to result in the formation of stigmata through which diapedesis occurred. Rabl, in 1893, suggested that the inter-endothelial lines, made visible with silver nitrate, were due to the formation of a silver proteinate.

The capillary circulation in the frog's mesentery has proven to be an excellent region for experimental studies on the rôle of the interendothelial cement (cf. 10). Artificial perfusates containing ash-free gelatin were used, the calcium content and pH of which were varied within viable limits. The gelatin was used since without it the Ringer's solution, even with the normal complement of calcium and

at the blood pH of 7.4, soon resulted in excessive edema indicating an abnormal permeability of the capillaries.

Changes in the physical state of the inter-endothelial cement could be detected by the adhesiveness to it of carbon particles suspended in the circulating perfusate. An increased sticking indicated a softening of the cement. Adhesiveness of India ink to the endothelium has generally been taken, e.g. by Landis, to indicate abnormal increase of permeability. Under normal conditions only occasional particles of carbon adhere to the vessel wall. When the acidity of the perfusate was increased or its calcium content was decreased there occurred an increased sticking and accumulation of the carbon particles especially along the inter-endothelial lines of the capillary wall. The softening of the cement was accompanied by a marked edema. The accumulation of the carbon along the endothelial lines indicated where the outward filtration of fluid was most pronounced. When excess calcium was present no such accumulation of carbon occurred along the inter-endothelial lines; moreover, no edema developed.

Under normal conditions the cement is being continually replaced. This was graphically demonstrated by applying a small quantity of 10 per cent silver nitrate with a micropipette to the surface of a capillary. The inter-endothelial lines immediately blackened and, as the blood flow continued, a blackened, amorphous material dropped off into the stream. The absence of any sign of leakiness or visible alteration in the flow indicated a continuous replacement of the cement as it was washed away. Micro-trauma invariably is accompanied by increased leakiness in the affected area. With more severe trauma a striking phenomenon is the appearance of transparent, glutinous masses exuding from the inner wall. These masses, made obvious by the carbon sticking to them, slough off and are carried downstream. The increased porosity of the vessel wall is indicated by a slowing of the blood stream as it approaches and passes by the affected region. In the region of sluggish flow the red cells become crowded together. After a few minutes the normal rate of flow returns with a recovery of the capillary wall to its former normal state. The reversibility of the phenomenon, from the development of a leaky state to eventual recovery, indicates a reconstitution of the cement which had been lost following the initial injury.

The suggestion is offered that an important rôle of the endothelial cell is the elaboration of an intercellular cement, the chemical stability and reactivity of which control the permeability of the blood capillary.

2. *Endocapillary Layer.* Another component of the capillary wall is a thin, non-cellular layer lining the inner surface of the endothelium. Its formation has been ascribed to the adsorption of a blood protein (13).

Danielli, in studies on perfused hind-legs of the frog, has offered indirect evidence for an adsorbed protein layer based on differences in the porosity of the capillary wall which occur with variations in the colloidal content of the perfusion fluid. He used a variety of colloids in isosmotic concentrations and found the order of relative efficiency in preventing edema, i.e., reducing capillary permeability to be: serum > acacia > ovalbumin > hemoglobin. He postulated that

certain blood proteins have a specific tendency to be adsorbed on the walls of the capillary pores and suggested that the particle size of the colloidal constituents play an important rôle. Such an adsorption mechanism would explain the findings of Krogh and Harrop (14) and of Drinker (15) that the edema caused by a perfusate of 3 per cent gum acacia can be markedly reduced by the addition of small amounts, e.g., 10 per cent, of blood serum.

Danielli also brought up the interesting point that certain basic proteins, e.g., clupein, added to serum, cause an increase in capillary permeability when the perfusate is kept at a constant pH. He suggested that the increased permeability is due to a displacement of the adsorbed plasma protein layer by the more highly surface-active clupein which forms a much thinner layer than does plasma protein.

In our investigations on the action of calcium and of pH on the inter-endothelial cement we were struck with the fact that perfusion fluids consisting of purely crystalloidal solutions, even with adequate calcium and proper pH, induced leakiness of the capillary wall with consequent edema although there was no indication of deficient inter-endothelial cement. This suggested that in the absence of a colloid the cement filter, although intact, lacks a physical component which is furnished by a colloid. The cement filter is porous enough to permit passage of colloidal material. A change in porosity occurs with the addition of slight amounts of a colloid which probably clogs the cement filter, thereby diminishing its porosity. Evidently, the adsorbed colloid constitutes the endocapillary layer. We used various colloids and found their relative efficiency in causing a recovery of a leaky capillary to be of the order: blood serum > bovine albumin > ash-free gelatin > gum acacia. This accords with similar findings of Danielli in relation to perfused blood vessels.

Direct microscopic evidence for such a layer has not been forthcoming. A near approach has been from observations of the frog's mesentery perfused with a crystalloid, frog-Ringer's solution containing Evans blue which is known to combine with blood albumin. During the perfusion thin strands and sheets of a faintly colored blue, translucent material we e seen sloughing off the inner surface of the capillary and being carried away in the stream. This sloughing continued for 10 to 15 minutes during which the mesentery became increasingly edematous. Changing the perfusate to one containing either 1 per cent gelatin or 10 to 20 per cent frog or fowl serum resulted in an appreciable reduction of the edema. When the perfusion fluid was changed back to the original crystalloidal solution with Evans blue there again appeared the sloughing off of faintly colored sheets of glutinous material.

3. *Pericapillary Sheath.* The existence of a supporting tissue investing blood capillaries has been suggested by several investigators, e.g., Heimberger (16), Benninghof (17), Midsumo (18) and Michels (19). Volterra (20a) made extended studies on this subject and has clearly demonstrated the presence of a closely fitting, investing layer of argentophil fibrils over the outer surface of the endothelium of capillaries. Such an investment is well exhibited in brain tissue.

In other tissues it appears to be continuous with argentophil fibrils surrounding glandular alveoles, muscle fibers, nerves, etc.

Volterra obtained variations in the amount of blood flowing through a given tissue to be in accordance with changes in pH and electrolytic content of the perfusate. He ascribed this to differences in action of the perfusate on the pericapillary sheath thereby conditioning the size of the vessels in the capillary bed. The significance of Volterra's conclusions lies in his conviction that variations in the amount of the circulating fluid may be referred to the supporting tissue without the necessity of assuming changes in the endothelium proper. This conclusion was not verified by looking for actual changes in the capillary structure.

In an earlier paper (20b) he describes, in cases of nephrosclerosis and hypertension, a transformation of the pericapillary sheath from the normal delicate layer containing finely anastomosing argyrophilic fibrils into a thick layer which may be hyalinized. Under such conditions he observed the presence of red blood cells permeating the wall and believed this to indicate an increased porosity of the vessel. More recently, Gorev and Smirnova-Zamkova (21) described similar changes, in cases of edema and hypertension, in a layer which they termed a hematoparenchymous barrier. The existence of a pericapillary sheath and its possible rôle in various pathological conditions makes a study of the normal properties of the sheath one of prime importance.

The presence of a relatively stiff layer of material against the outer surface of the capillary endothelium is evident when a white blood cell is undergoing diapedesis. The extruded portion of the diapedesing cell never moves directly away from the wall but spreads over the outer surface of the endothelium and remains for some time pressed closely between it and the pericapillary sheath. The cell finally works its way through the interstices of the sheath to move more freely in the less resistant regions of the connective tissue matrix. The relation of this sheath to the connective tissue is suggested by the fact that agents which affect the connective tissue matrix appear to exert a similar action on this pericapillary sheath. We have made intravenous injections of the mucolytic enzyme, hyaluronidase,[3] and also have applied solutions of it with micropipettes on the surface of capillaries in the mesentery of the frog. In both cases there was no evidence of increased stickiness of the capillary wall or of any increase in permeability. A significant feature which did occur was the abrupt development of microscopic petechial hemorrhages. These occur through spots in the wall weakened by the softening of the supporting connective tissue sheath and, therefore, easily ruptured by the internal blood pressure. We regard hyaluronidase as a factor in accentuating capillary fragility rather than in inducing direct changes in capillary permeability (22).

Moreover, we found that bacterial toxins[4] (Shiga exotoxin, Clostridium

[3] The hyaluronidase was prepared from testicular extracts and supplied through the courtesy of Dr. Karl Meyer of the College of Physicians and Surgeons, Columbia University, N. Y.

[4] These bacterial toxins were supplied through the courtesy of Dr. René Dubos of the Rockefeller Institute and Dr Colin MacLeod of New York University College of Medicine.

welchii, C. perfringens and C. septicum), injected into the blood stream of rats, cats and dogs, caused the production, several hours later, of petechial hemorrhages among the mesenteric and omental capillaries with no evident dissolution of the inter-endothelial cement. These bacterial toxins are known (23) to contain a hyaluronidase component to which the increased fragility of the capillaries can be attributed. We therefore regard Duran-Reynal's (cf. 22) postulate that hyaluronidase regulates capillary permeability as being concerned with an increase in fragility of the capillary wall rather than with the more readily reversible phenomenon of changes in capillary permeability.

4. *Micromanipulative Studies on the Capillary Wall.* The structural make-up of the normal capillary wall is so delicate that mechanical handling even on a micrurgical scale invariably produces changes in its permeability characteristics. This has already been indicated by Landis (cf. 11). The reversibility of such an injury depends upon the intensity of the change produced. Merely moving the tip of a microneedle back and forth in the close vicinity of a capillary will inflict an evanescent injury sufficient to cause an outflow of fluid. This is made evident by an approach, to the injured spot, of red cells and an outward leakage of dye. In mammals there also occurs a temporary sticking of platelets to the inner surface of the capillary wall. A more drastic reaction, but one which is still reversible, can be occasioned by gently rubbing the surface of the capillary with the blunt tip of a microneedle. This induces not only a sticking of platelets and of leucocytes but also a rapid escape of dyes of relatively large colloidal dimensions, such as Evans blue. The leakiness of the wall may become so high that the red cells are diverted from their forward course and applied against the inner surface of the vessel. With still more severe injury but insufficient to cause actual rupture, red cells may be passed through minute openings to the outside. The various effects, grading from a slightly increased porosity to a pronounced leakiness, are similar to those observed when perfusion fluids are used which are deficient in calcium or are more acid than normal. Similar effects are obtained when the adjacent tissue is acidified. Evidence of this has been obtained (24) by depositing a microdrop of olive oil containing glacial acetic acid in the vicinity of a capillary.

All the procedures, viz., microtrauma, use of calcium-deficient perfusates, calcium-containing, acid perfusates, or local acidification, act by softening (made evident by carbon sticking) and dissolving the inter-endothelial cement. The changes in the cement caused by the microtrauma are possibly due to a local production of an "acid of injury". An additional factor is the contracting reaction of individual endothelial cells. Contiguous cells can be made to withdraw from one another sufficiently to loosen the cement between them. An extreme example of the latter type of change is the shrinkage reaction of the endothelial cells to hypertonic sucrose or mannitol solutions locally applied with a micropipette. The cells retract from one another to the extent that spaces between them become studded with dumb-bell shaped red cells partly extruded through the vessel wall. As the hypertonicity passes off the endothelial cells return to their normal expanded condition and close the spaces

between them. The red cells, caught in the wall, then either move out, fall back into the vessel, or become pinched in two (25).

SUMMATION ON SIGNIFICANCE OF ARCHITECTURE OF CAPILLARY WALL. The preceding analysis of the constitution of the capillary wall presents a structural basis concerned with the interchange of material between blood and tissue. The inter-endothelial cement serves as the basic framework, the porous interstices of which vary with changes in the electrolytic balance of the medium. The endocapillary lining appears to be an adsorbed layer of a blood protein which penetrates the interstices of the cement and serves further to regulate the pore size of the cement filter. The third constituent is the pericapillary sheath which appears to be a condensation of connective tissue serving to give mechanical support to the blood capillary. It is probably sufficiently porous to allow free passage for the fluid part of the blood.

Each of the above three components can be affected separately by a variety of conditions so as to induce changes in the permeability of the wall. Hence it is difficult to assume a single permeability factor, such as a permeability hormone or vitamin, to account for the overall maintenance of fluid balance between blood and tissue.

It is to be noted that variations in the permeability of the capillary can be accounted for by changes in the non-living components of the capillary wall. An important function of the endothelial cells appears to be the elaboration of the cement. In extreme cases, since they are irritable, the cells can contract or expand and in this way affect the porosity of the wall by changing the physical characteristics of the cement filter.

TOPOGRAPHY OF THE CAPILLARY BED. Since capillary permeability is intimately related to hemodynamic aspects of the peripheral circulation, it is necessary to review briefly the factors concerned in the distribution of blood through the capillary bed. It is unfortunate that an understanding of the actual anatomical topography of the capillary bed has been generally overlooked in attempts to elucidate its functional activities. The tendency has been to consider the capillaries as simple tubes with the only anatomical differentiation of classifying them as arterial or venous.

Topographical studies of the capillary bed have shown that the vascular components may be of different sorts which are precisely arranged into organized units of structure and function (26). In some tissues the capillaries have no discernible organization. This type is found in tissues which maintain a fairly constant level of flow volume. In other tissues, where nutritive demands vary with changes in functional activity, a well defined pattern exists, e.g., in the muscular system and in the gastro-intestinal tract. During periods when the tissue is inactive the flow is restricted to preferential channels, while during periods of increased activity the flow becomes widespread throughout the capillary network.

The preferential vessels have been termed thoroughfare or a-v channels. The proximal portion of these channels, termed metarterioles, together with their precapillary sphincteric offshoots, are muscular and spontaneously undergo

periodic changes in caliber. This type of movement has been termed vasomotion, a slow intermittency of partial relaxation and constriction at intervals of about 30 seconds to 3 minutes. The vasomotion serves as a mechanism for restricting the blood flow to the thoroughfare channels or permitting it to become more widespread through the capillary bed. It is the mechanism which locally adjusts the peripheral flow to the needs of the tissue. The precapillary sphincteric offshoots lead into an inter-anastomosing system of true capillaries (devoid of muscular elements) which constitutes the bulk of the bed. The capillaries rejoin the distal continuation of the thoroughfare channels through inflowing tributaries.

The hydrostatic pressure in any of the true capillaries throughout the bed is extremely variable. It depends from moment to moment on the intermittent state of the precapillary sphincters and on the efficiency of the venous outflow. On the other hand, the thoroughfare channels maintain a proportionally constant pressure relationship between their arteriolar and venous ends. This condition depends largely upon the pressure in the arterioles leading into them. The venous outflow from the bed is determined by the rate and magnitude of the flow through the thoroughfare channel.

Thus, the thoroughfare channel is the basic structural and functional component of the capillary beds. The capillary network is accessory to it and exhibits great variability in its rôle in fluid exchange. This variability was noted by Landis and is reflected in the high degree of scattering exhibited in recorded measurements on intracapillary pressure (Landis (27)). Such measurements would be of greater significance and more consistent if they were made on the arterial and venous segments of the thoroughfare channels.

The relatively low hydrostatic pressure in the true capillaries throughout the bed is accomplished by the resistance offered through the sharp, backward twisting of the junctional portions of the outflowing precapillary offshoots, together with the relatively wide postcapillaries which lead into the distal portion of the thoroughfare channel. The sphincteric action of the precapillary musculature is a further factor in sharply lowering the arterial pressure as the blood enters the true capillaries.

The flow through the thoroughfare channels is usually distinctly more rapid than elsewhere in the bed, and the fall in hydrostatic pressure from their arteriolar to their venous ends must be of lesser magnitude than that which occurs in the side branches. As a result, the channel is to be considered as the predominating region for outward filtration while inward filtration would be a characteristic function of the true capillaries. The occurrence of outward and inward filtration at different sites makes it unnecessary to assume that the movement of fluid and of suspended materials should occur simultaneously in opposite directions through the wall of the same capillary vessel.

Vasomotion and its influence on fluid exchange[5]. Vasomotion is the

[5] This and other sections constitute an amplification and an analysis of accumulated observations from a group study on the capillary circulation in secondary shock by the authors with E. L. Chambers, G. H. A. Clowes, G. W. Duncan, C. G. Grand, C. Hyman, M. J. Kopac,

term applied to the spontaneously occurring periodic relaxation and constriction of the thoroughfare channel and of its precapillaries (cf. 10). The dilator and constrictor phases of the vasomotion alternate irregularly from minute to minute, and the duration of one or the other phase varies under different physiological conditions. The vasomotion of the muscular, proximal portion (the metarteriole) of the thoroughfare channels is wavelike and exerts a milking action on the inflowing postcapillaries which lead into the distal portion of the channel. Normally, the blood-flow through the channel is fairly constant as a result of which its hydrostatic pressure should remain relatively high. The vasomotion of the precapillary offshoots, manifested by an opening and closing of their sphincters, conditions the flow through the true capillaries among which it occasions alternate periods of varying hydrostatic pressure. The relative duration of these periods varies greatly under different conditions. The fluid exchange in the capillary bed is thus greatly influenced by the character of the vasomotion. This delicately balanced activity depends upon a nervous factor, which is vaso-excitor, and on humoral factors which may be either vaso-excitor or vasodepressor. The tissue origin of some of the humoral factors offers a means for the vasomotion to act as a mechanism for locally regulating the distribution of the blood and, thereby, the extent and duration of inward and of outward filtration in the capillary bed.

The direct action of the vasomotion affects the fluid exchange by determining the flow of blood to be either by way of the thoroughfare channels or of the true capillaries. The fluid exchange is also affected indirectly by the influence which the distribution has on the rate of venous outflow from the capillary bed. For example, when vasomotion is deficient or absent, the precapillaries remain open and the propelling force of the pressure transmitted by the arterioles becomes dissipated through the numerous capillaries of the bed. This spread of flow causes a slowing in the rate of overall movement through both capillaries and thoroughfare channels with the result that blood tends to accumulate in the collecting venules. This induces sufficient back pressure to favor an overall outward filtration. On the other hand, when vasomotion is active, the flow through the capillary bed is increasingly restricted to the preferential channels so that the propelling force of the arterial pressure through them is sufficient to enhance a rapid flow in the collecting venules. The rapid venous flow favors drainage into the venules from the true capillaries in which inward filtration is thereby accentuated.

We may recapitulate the influence of vasomotion on the effectiveness of the hydrostatic and of the colloid osmotic pressures in the different regions of the capillary bed. Variations in vasomotion are of two sorts: first, alterations in the rate of the intermittent caliber changes of the thoroughfare channel and,

---

M. E. Krahl, R. E. Lee and B. E. Lowenstein, published and being published in the Am. J. Anat., Am. J. Physiol., Annals of Surgery, Proc. Exper. Biol. and Med., and Surgery, Gynec. and Obstetrics. Much of the work was presented in the periodic reports sent to the Committee on Medical Research of the O. S. R. D. and to the subcommittee on Aviation Medicine of the National Research Council during the years of 1941 to 1945.

second, alterations in the relative duration of the constrictor and the dilator phases of the precapillary sphincters. The most effective hydrostatic pressure occurs in the thoroughfare channel along the length of which the onward flow is fairly constant and rapid. Moreover, the rate of vasomotion of the channel determines the rate of venous outflow from the bed by its influence on the inflow from the postcapillaries which lead into its venous end.

The colloid osmotic pressure should be most effective in the network of true capillaries for absorbing fluid from the tissue. However, its effectiveness is variable and depends upon shifts in the hydrostatic pressure conditioned by the relative duration of the successive constrictor and dilator phases of the precapillary sphincters. The consequent periods of flow and of no flow in the true capillaries constitute a mechanism whereby the colloid osmotic pressure is periodically greater than the hydrostatic pressure. The existence of varying pressure conditions was inferred by McMaster (28) who demonstrated the intermittency of fluid uptake in the subcutaneous tissue by inserting a micropipette into the skin.

Finally, the hydrostatic pressure in the thoroughfare channels also depends on the vasomotor activity of the terminal arterioles outside the capillary bed. This activity of the terminal arterioles tends to maintain the hydrostatic pressure in the thoroughfare channels at a constant level despite fluctuations in the systemic blood pressure. The arterioles accomplish this by undergoing constriction when the arterial pressure rises, and by dilating when the pressure falls. Under certain conditions the arterioles can be affected, e.g., after an intravenous injection of fever-producing toxins (unpublished data). The arterioles then become dilated while no change occurs in the normal rhythm of the vasomotion of the thoroughfare channels. This phenomenon is accompanied by hemoconcentration made evident by close packing of blood cells in the collecting venules. The following mechanism is suggested for this condition. The dilated arterioles permit the blood to enter the capillary bed under an increased head of pressure. At the same time the maintenance of the normal periodicity of the vasomotion favors the flow through the thoroughfare channels. This produces an excessive outward filtration as a result of the increased pressure in the channel. On the other hand, the true capillaries maintain their normal intermittency with no corresponding increase of inward filtration. The result is that the overall balance favors outward filtration with consequent hemoconcentration.

It is also possible to have hemodilution with no change in the vasomotion. This occurs when the head of pressure entering the capillary bed is significantly reduced. Such a situation occurs immediately after the onset of bleeding in acute hemorrhage. The larger arteries and arterioles undergo widespread vasoconstriction to compensate for the suddenly reduced blood volume. The constriction also occurs in the terminal arterioles so that the effective hydrostatic pressure in the thoroughfare channels is lowered with a consequent reduction of outward filtration. However, vasomotion remains active with good venous outflow so that inward filtration is maintained at a normal rate in the true

capillaries. This shifts the balance in favor of inward filtration, resulting in hemodilution.

The above phenomena illustrate imbalances in fluid-exchange in the capillary bed with no disturbance in vasomotion, the imbalance resulting only from changes in the vasomotor activity of the terminal arterioles. There are also instances in which a change in the intensity of vasomotion alters the balance of fluid-exchange. For example, when the loss of blood during acute hemorrhage exceeds two per cent of the body weight, the vasomotion becomes accelerated with a still further accentuation of hemodilution. This can be explained as follows. As the frequency of the vasomotion increases, the duration of the constrictor phases becomes prolonged relative to the dilator phases. Concomitant with the long constrictor phases which reduce transmission of pressure into the side branches there is a progressively longer period during which the hydrostatic pressure in the true capillaries remains extremely low. As a consequence, inward filtration and tissue dehydration become still further accentuated. An additional aid in this direction is the increased frequency of the precapillary sphincteric contractions, producing a pronounced intermittency of the flow through the capillary bed. Moreover, the wavelike milking action of the thoroughfare channel facilitates drainage of the post-capillaries into the venous circulation.

We have been considering instances in which either the vasomotion or the state of contraction of the feeding arterioles can be affected separately. When both are suspended simultaneously, edema results. A cessation of vasomotion together with a dilated state of the feeding arteriole flushes the entire bed, resulting in excessive outward with little or no inward filtration.

In summary, deficiency or absence of the vasomotion shifts the balance of fluid exchange in favor of outward filtration while enhanced vasomotion shifts the balance to inward filtration. Thus, the hemodynamic relations of vasomotion and the forces set up by movement of fluid through one or another of the vascular components of the bed must be regarded as of major significance in affecting fluid exchange between the blood and tissue. It follows that some types of edema are the result of disturbances in the vasomotion mechanism and do not necessarily involve alterations in the permeability of the capillary wall.

A list of the factors which influence vasomotion is of interest in this discussion. Enhanced vasomotion occurs with acute hemorrhage, sympathetic stimulation, intravenous administration of adrenalin, of angiotonin or of adrenal cortical extract. Diminished vasomotion occurs with a rise in temperature within viable limits of 37.5 to 41°C, decrease in temperature, direct trauma, increased vital activity (muscular exercise, secretory activity, etc.), or elaboration of vasodepressor principles following prolonged anoxia. Many agents which have been found to increase capillary circulation do not necessarily depress vasomotion. Among these have been found histamine, adenylic acid, adenosine, kallikrein and acetylcholine (unpublished data).

Vasomotion and lymph flow. There is an intimate relation between vasomotion and the rate of lymph flow in the terminal lymphatic channels. This was studied in the mesentery of the cat, dog, mouse and rat During the resting state of a tissue the vasomotion is active to the extent of producing relative ischemia, and little or no flow of lymph can be discerned in the terminal lymphatics. The absence of an active flow throughout the bed makes the large surface furnished by the true capillaries available for inward filtration and leaves no excess in the tissues.

On the other hand, during the active state of a tissue the vasomotion becomes diminished so that an overall, hyperemic flow is favored. This is accompanied by an appreciably increased flow of lymph in the terminal lymphatics. The prolonged periods during which the precapillary sphincters remain dilated favor outward filtration in the true capillaries Edema is prevented only by an absorption of fluid into the lymphatic capillaries

Gradient of permeability. Starling's (29) stimulating concept of fluid transfer between blood and tissue is undoubtedly true as an overall aspect of the phenomenon. According to this concept outward filtration occurs where the hydrostatic pressure is greater than the colloid osmotic pressure of the blood while inward filtration occurs where the conditions of pressure are reversed. Difficulties arise when one attempts to allocate specific regions in the capillary bed where outward and inward movements of fluid occur. According to Starling's postulate, the wall of the capillaries throughout a given tissue possesses uniform characteristics of permeability. Fluid transfer under physiological conditions would then depend solely on pressure relationships of hydrostatic and collloid forces.

Landis (27, 30, 31, 32, 33) in his well-known micromanipulative determinations of capillary blood pressure, found considerable variation in the values obtained. By averaging the pressure readings, he found that those from the arterial side of the bed tended to be significantly higher, and that those from the venous side of the bed were lower than the known colloid osmotic pressure of the blood. He concluded that outward filtration of fluid and its absorption through the wall of the capillaries is a function of the observed intracapillary pressure. Confirmatory findings were obtained by Gordon-Königes and Ottó (34) from their micromanipulative experiments on the capillaries of the intestinal villi of the cat. They found that the flow in the lacteal of the villus was correspondingly greater the higher the capillary blood pressure.

Landis takes the view that the preponderantly higher blood pressures obtained by him among the capillaries in the arterial regions of the bed in the frog's mesentery, indicate that outward filtration, on the average, occurs in the arterial region of the bed while absorption of fluid from the tissues occurs in the venous capillaries where the average hydrostatic pressure tends to be lower than the colloid osmotic pressure of the blood. He obtained evidence in support of this from the ingenious experiment of placing the blunt end of a micro-rod across a flowing capillary so as to occlude its lumen locally and obstruct the flow. He obtained values for the pressure in the arterial end of the capillary proximal

to the obstruction indicating that it had assumed the value of that of the vessels which had been flowing into it, while the pressure distal to the obstruction fell to a minimum and acquired a value corresponding to the vessel flowing from it. The crowding together or the moving apart of the corpuscles in the two segments of the capillary, taken by him to indicate outward or inward filtration, was consistent with the degree of pressure he recorded in relation to the known colloid osmotic pressure of the blood. The procedure of occluding the capillary with a micro-rod requires very careful technique because of the extraordinary ease with which damage may be inflicted, causing an abnormal increase in porosity of the wall in the obstructed region. Furthermore, it is obvious that the pressure conditions induced by the obstructing needle do not necessarily prevail in the same capillary when the obstructing needle is not there. We have obtained similar results with capillaries irrespective of the position of the capillary in the capillary bed provided the capillary exhibits a good flow prior to the placement of the obstructing needle.

In his 1926 paper (cf. 33) Landis distinguished between the capillaries which are in more or less direct relation with the finest arterioles and the venous capillaries which merge to form the venules. He noted particularly that the capillaries, where they spring from the arterioles, are narrow and at times exhibit "slit-like openings". He remarked on the conditions at the arteriole-capillary junction as being related to his observed "rather sharp reduction in the pulse pressure on passing from the arteriole to the capillary". The junctional regions of these capillaries are doubtless the vessels we have termed the precapillaries, possessing muscular sphincters, and which we have described as springing both from terminal arterioles and from the periodically constricting proximal portions (metarterioles) of the thoroughfare channels.

The question now arises as to the regional distribution of the vessels in the capillary bed which are concerned with the outward and those with the inward transport of fluid across their walls. Rous, Gilding and Smith (35), arguing from the more rapid escape, at the venous end of the capillary bed, of dyes introduced into the blood stream, claimed that the porosity of the vessel at the venous was greater than at the arterial end and that outward filtration also occurred there.

Landis (36) dismisses the possibility that the diffusion of dyes is always a dependable measure of the direction and rate of fluid transport since diffusion may conceivably occur in the opposite direction to fluid flow. Danielli, in support of the Starling theory, has offered the hypothetical explanation that the observed accumulation at the venous end is due to the reabsorption of fluid leaving behind whatever dye has escaped.

The suggestion is made here that the introduction of a concept regarding the capillary bed as possessing a pattern of specialized units of function built about the thoroughfare channels may bring into harmony the seemingly divergent points of view held by Landis and by Rous and his co-workers.

According to our view the thoroughfare channel is pre-eminently the site for outward filtration while the true capillaries, which branch off from the

channels (as outflowing vessels at the proximal and as inflowing at the distal end of the channel), serve largely for inward filtration especially during the resting stages of the tissue involved.

Landis himself commented on the great variability of the values in his recorded capillary pressure readings. He states that the pressures differed not only in different capillaries coming from the same arteriole but also in the same capillary from moment to moment. On the average, the more rapid the flow in a given capillary the higher was the recorded pressure. This variability may be accounted for, at least in part, by the periodic opening and closing of the precapillary sphincters. The conditions affecting this activity in ischemia and hyperemia are discussed elsewhere in this article.

The findings of Smith and Dick (37a) are of interest in this regard. They introduced into the circulation a continuous stream of glucose immediately after an intravenous injection of Chicago blue 6B, a poorly diffusible colloidal dye. They found a great increase in the plasma volume, indicating passage of fluid from the tissues to the blood. Despite the evident dehydration of the tissues, they observed that the accumulation of the dye outside the capillary blood vessels occurred in the same venous regions as described by Smith with Rous and others. They concluded that the movement of fluid from the tissues to the blood does not essentially alter the gradient of vascular permeability. Moreover, their findings have been interpreted as meaning that outward diffusion of dye and inward filtration or absorption of fluid can occur simultaneously in the same vessels.

This raises the issue regarding diffusion of a dye against a current of moving fluid. In our observations the diffusion of dye outward through the wall of the thoroughfare channel is too rapid to permit any other assumption than that the dye is being carried with the stream of the fluid in which it is suspended. In the words of McMaster and his associates (37b) concerning dyes in aqueous solution "where water goes they may be expected to go, unless their molecules are large enough to be held back seloctively by the membrane".

We may now consider the alternative that the influx of fluid occurs in a region of the capillary bed other than that from which the dye was escaping. According to our findings, the inflow of fluid which makes for hemodilution occurs in the true capillaries where the hydrostatic force is generally insufficient to drive fluid out. In the case of Smith and Dick's experiments the hypertonicity of the sugar solution would increase the effectiveness of the colloid osmotic pressure of the blood in absorbing fluid into the true capillaries. At the same time the escape of dye would continue along the thoroughfare channels which are increasingly porous along their distal portions and which evidently maintain sufficient hydrostatic pressure to overcome the effect of the hypertonicity induced by the sugar solution.

The accompanying diagram (fig. 1) presents three concepts regarding fluid movement in the capillary bed.

The capillary bed is represented as a tube passing from *a*, the artery, to *v*, the vein. The arrows show the direction of blood flow and of fluid movement across the capillary wall.

X presents the concept of Rous and his associates in which the arrows only apply to the preponderancy of dye-escape at the venous end.

Y presents Starling's concept as demonstrated and amplified by Landis from his capillary blood-pressure determinations. In this, the fluid movement is predominatingly outward at the arterial end and inward at the venous end.

Our concept combines *X* and *Y*. *X* represents the thoroughfare channels in which outward filtration occurs together with the phenomenon of vasomotion which intermittently opens and closes the precapillary sphincteric offshoots. The latter lead into *Y*, the true capillaries, in which the inward and outward filtration of fluid depends on continually occurring variations in the capillary blood pressure produced by the vasomotion of the sphincters. Toward the venous end the variability of pressure shifts to favor inward filtration or absorption.

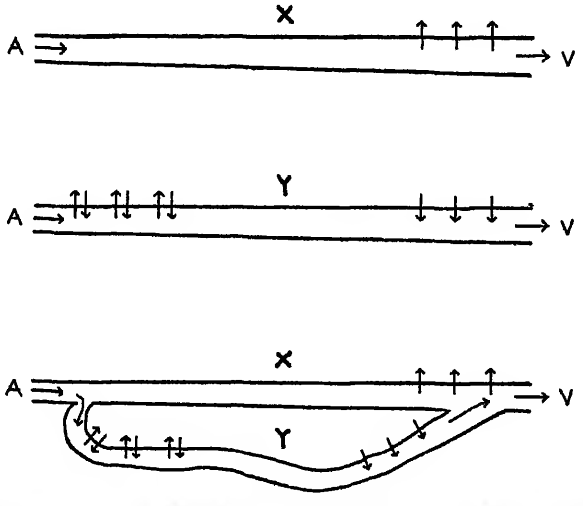


Fig. 1. Presenting three concepts regarding fluid movement in capillary bed.

The post-capillaries widen as they join the distal portions of the thoroughfare channels at an acute angle in such a manner that the blood in them is drawn into the channel analogous to the way that a vacuum jet operates.

Variations from the third sketch, *X* and *Y*, do occur, e.g., the precapillaries frequently lead off directly from terminal arterioles and the post-capillaries may lead directly into typical venules. However, the functional feature is the same, i.e., the intermittency of action of the precapillary sphincters and the peculiar arrangement of the post-capillaries. Fluctuations of pressure in the post-capillaries may occur from the occasional development of a significant venous back-pressure. Such fluctuations may change the direction of movement across the capillary wall.

The gradient of permeability in the thoroughfare channel seems to represent a gradation in the limiting dimensions of the pores along the wall of the vessel. This is indicated by introducing dyes of graded particle size into the circulation and noting their escape along the channel. The results have warranted the

conclusion that there is a gradual increase in the relative number of the larger pores toward the end where the blood is most venous.

The experiments of experimentally re-routing the venous blood indicate that the size and number of pores is a variable feature (38).

The direction of the gradient can be upset experimentally. By compression of appropriate channels with a microneedle, the flow through the capillary bed was reversed. After a short period of time a new gradient was established, an increasing porosity now becoming evident toward the original arterial side of the capillary bed. Thus, when blood from the venous vessels was re-routed back through the thoroughfare channel, a marked increase in the porosity of the vessel wall occurred, colloidal dyes such as T-1824, to which only the venous end is normally permeable, now passing out along its entire length. The gradient, therefore, appears to be a result of some factor in venous blood and not due to permanent, structural differences in the vessel wall.

A gradient of porosity is also of significance in the passage of water through the capillary membrane. The diameter of the pore is of prime importance since any increase in diameter affects the magnitude of the filtration to the second power of the diameter. Therefore, changes in hydrostatic pressure should cause a proportionally greater increase in outward filtration of water on the venous side where the pores are larger than on the arterial side. We find the gradient of porosity to be exclusively along the thoroughfare channels through which, from their arterial to their venous ends, the flow of blood is relatively rapid and along which outward filtration occurs. The existence of a gradient of porosity in the capillary bed must be taken into account in considering the effect of hydrostatic pressure on the outward filtration of fluid.

Summation of significance of vasomotion. Attempts in the past have been made to offer precise quantitative support for Starling's hypothesis of fluid exchange in the capillary bed. Landis has presented calculations indicating a precise counterbalance between capillary hydrostatic and colloid osmotic pressure, the loss of fluid on the arterial side being compensated for by an equivalent intake of fluid on the venous side.

In the present review, the basic concept of Starling's hypothesis remains unchallenged. However, the balance of fluid exchange is considered to be maintained not so much by a direct interplay of hydrostatic versus colloidal pressures as by delicate vasomotor adjustments which intermittently increase or decrease the surface area over which an effective hydrostatic pressure produces outward filtration.

Inward filtration is accomplished primarily by osmotic uptake. The effectiveness of this mechanism, however, is directly dependent upon and is reinforced by the action of hemodynamic forces maintaining an adequate venous outflow from the capillary bed. Keys and co-workers (39) and Ralli and her co-workers (40) found that edema can occur with no concomitant reduction of the plasma proteins in the blood. Such data indicating the relative ineffectiveness of colloid osmotic pressure for preventing fluid loss from the capillary circulation emphasize the importance of factors other than the interplay of hydrostatic

vs. osmotic pressure in maintaining normal fluid exchange between the blood and the tissues. Their findings may be explained either by the assumption that the permeability of the capillary wall is so affected as to cause the loss of whole plasma, or that the loss of non-protein fluid in one region, e.g., the abdominal cavity or dependent parts of the extremities, is counterbalanced by an uptake of similar fluid from another region, e.g., skeletal muscle.

NERVOUS CONTROL OF CAPILLARY PERMEABILITY. It was claimed by Engel (41) that the sympathetic system induces an increase in capillary permeability, a view opposite to a previous one of Gellhorn (42) who presented evidence that sympathectomy induces an increase. Engel perfused the vessels of the knee-joint of a cat with fuchsin S in Ringer's solution and found that, although sympathectomy induced marked vasodilatation, the concentration of the dye in the knee-joint fluid was lowered. He ascribed this to a decreased permeability resulting from the absence of sympathetic nerve control. Danielli (cf. 9) discounted Engel's findings as indicating permeability changes and ascribed them to a lowered capillary pressure resulting from vasodilatation. Engel's results might more logically be explained by the opening of cross anastomoses (AVA) between the larger vessels so that the capillary bed in the particular region he was observing was bypassed by the active circulation.

Most authors favor Gellhorn's original thesis and have concluded that sympathectomy produces an increase in loss of fluid from the capillaries. This would be expected on the basis of the resulting hyperemia and increased blood flow. However, this does not necessarily signify any specific effect on endothelial permeability since the resulting vasodilatation and hyperemia should cause an increase in hydrostatic pressure in the capillary bed. This, together with suspension of vasomotion, should greatly favor outward filtration.

EFFECT OF AGENTS ON CAPILLARY PERMEABILITY. The literature contains many references to the action of a large number of biological agents in affecting the permeability of the capillary wall. A list taken from the more recent literature is presented in the accompanying table. Only those investigations are listed which deal with findings purporting to be directly related to the capillaries.

A commonly used technique (43) is that of observing the loss of dye from the blood into the skin during a local inflammatory reaction and noting the effect of various locally introduced agents in altering the amount of dye escape. Such studies shed little or no light on the mechanism of capillary permeability. Agents, such as desoxycorticosterone acetate, estrogens, and cortical extracts, which have been found to reduce the accumulation of trypan blue in an inflamed area (see table) do not necessarily produce their effect by acting on the capillary wall proper. For example, Rigdon (55) who used this method extensively has questioned the significance of the dye accumulation since he found that the stainability of the affected tissue is involved. The inflammatory reaction is a complex one, involving such diverse effects as vascular dilatation, endothelial damage, changes in the connective tissue matrix, blockage of lymphatic flow, stainability of the tissue, etc. The value of such data in clarifying capillary permeability phenomena is open to question.

TABLE 1

Agents affecting capillary permeability

AGENT	AUTHOR	EFFECT	CRITERION
Adrenal cortical secretions	Swingle et al. (43)	Maintenance of capillary tone	Adrenalectomy: increased permeability and capillary stasis
Adrenal cortical secretions	Cope et al. (44)	Decreased capillary permeability	Adrenalectomy: increased blood protein in lymph
Adrenal cortical extract DCA	Menkin (45)	Decreased capillary permeability	Leukotaxin effect of dye accumulation in skin was decreased
Adrenal cortical steroids Cortico sterone	Freed and Lindner (46)	Decreased capillary permeability	Same as above
Adrenal cortical extract	Shleser and Freed (47)	Decreased capillary permeability	Peptone effect of dye accumulation in skin was decreased
Adrenal cortical extract 11-desoxycorticosterone	Hyman and Chambers (48)	Reduced edema formation	Decreased weight of perfused hind-limbs of frog
DCA Cortin	Graham (49)	Decreased capillary permeability	Dye leakage by carbon arc irradiation of skin decreased
DCA	Fine and Fischmann (50)	No effect on capillary permeability	Rate of appearance of dye in tissues, disappearance of wheals; effect on albuminuria
DCA	Swingle and Remington (51)	Decreased capillary permeability	Increased retention of transfused serum in adrenalectomized dogs
Estrogens	Hechter et al. (52)	Increased capillary permeability in uterus and vagina	Increased concentration of dye in tissues
Alpha estradiol benzoate	Rigdon and Chrisman (53)	No effect	Xylol effect of local accumulation of dye in skin unchanged
Vitamin D	Silver et al. (54)	Decreased permeability with high doses	Decrease in disappearance rate of T-1824 from blood

TABLE 1—*Concluded*

AGENT	AUTHOR	EFFECT	CRITERION
Tissue extract	Rigdon (55)	Increased capillary permeability	Local injection induces accumulation of dye in skin
Leukotaxin	Menkin (56)	Increased capillary permeability	Local injection induces accumulation of dye in skin
Spreading-factor	Duran-Reynals (57)	Increased capillary permeability	Local injection induces accumulation of dye in skin
Histamine	Roche e Silva and Dragstedt (58)	Increased capillary permeability	Local injection induces accumulation of dye in skin
Histamine	McCarrell and Drinker (59)	Increased capillary permeability	Increased blood protein content of lymph
Histamine	Stead and Warren (60)	Increased capillary permeability	Decreased protein in venous blood
Histamine	Hechter (61)	Increased capillary permeability	Local dye accumulation unaffected by adrenalectomy
Acetylcholine	Gordon-Königes and Ottó (34)	Increased capillary permeability	Increase of lymph in lacteal of villus following dilatation of arterioles and increased capillary pressure

Other studies have used the appearance in the tissue of agents, the concentration of which can be measured, as an indication of permeability changes. Lange (62) measured skin fluorescence by a photometer after intravenous injection of sodium fluorescin. He found a decreased rate of dye accumulation in the skin under conditions such as myxedema. This dye is freely diffusible through the normal capillary wall. Evidence of increased fluorescence in the skin under experimental or pathological conditions may merely indicate an increased blood flow through the tissue with no actual increase in the permeability of the capillary wall to the dye.

A similar criticism can be made of experiments on the permeability of radioactive ions. Experiments with radioactive proteins (Fine and Seligman, 63; Cope and Moore, 64) should be of greater significance since proteins are normally retained by the capillary wall. The experiments of Fine and Seligman, which were performed on dogs in irreversible shock, indicated no outward leakage of

the radioactive proteins. On the other hand, Cope and Moore found an increased loss of radioactive colloids into the lymph following burns.

Of the various biological agents investigated, histamine and adrenal cortical extracts are of special interest. It has been claimed that histamine increases capillary permeability while the cortical extracts reduce it.

Histamine. The pioneer work of Lewis (65) and of Lewis and Grant (66) indicate that histamine-like substances produce a triple vascular response terminating in the marked exudation of protein-containing fluid from the capillaries into the tissues. Histamine, introduced into the skin, produces both arterial dilatation and damage to the capillary endothelium. In our experiments we found that histamine, introduced either locally into the mesentery with a micropipette or systemically by intravenous injection in concentrations not affecting arteriolar caliber, had no observable permeability effects (judged by dye exudation, stasis, etc.). Application with a micropipette in concentrations which produce dilatation of the arterioles also had no observable effects. Permeability changes were obtained only when concentrations were used high enough to produce actual endothelial damage, i.e., extravascular loss of normally retained Evans blue, sticking of carbon, swelling of individual endothelial cells and an abnormal disten ion of the capillary wall. To what extent histamine plays a rôle in normal permeability changes is still open to question.

Other experiments indicating a permeability-increasing effect of histamine have relied upon the loss of materials from the blood, e.g., proteins, which are normally retained within the capillary. This was detected by analysing the lymph fluid, the venous blood from affected regions or the fluid in various body cavities. Obviously, damage to the endothelium, such as occurs with excessive concentrations of histamine, will produce abnormal leakage not only of plasma but of blood cells.

In many cases the extravascular loss of protein can be explained equally well by vascular mechanisms other than increased capillary permeability. In this regard two techniques can be considered, that used by McCarrell and Drinker (cf. 59) and that by Stead and Warren (cf. 60). McCarrell and Drinker used changes in the constitution of the lymph of a particular region as an index of blood capillary permeability. They found that histamine shock produced an increase of protein in the lymph. Such an occurrence has a possible explanation from the findings of E. R. Clark (67). Clark noted that the application of pressure or increased temperature caused the passage of whole blood through the wall of venules directly into contiguous terminal lymphatics. This might be sufficient to explain the increase in protein content of the lymph without postulating an increase in capillary permeability *per se*. Stead and Warren determined differences in the contents of arterial and venous blood from affected areas. They noted that an injection of histamine (0.5 mgm.) into the brachial artery caused a decrease of the protein content of the venous blood of the same arm. This finding may be interpreted as an effect of the peculiar type of vascular reaction caused by histamine. Histamine produces arteriolar dilatation and

increases the blood flow through the capillary bed but it does not abolish vasomotion of the precapillary sphincters. This condition produces an excess of blood flow through the thoroughfare channels accentuating outward filtration all along their course. Normally, only a small amount of protein escapes from the venous portion of the channel. However, the increased hydrostatic pressure produced by histamine could be expected to lead to a greater leakage of protein in this region. Concomitantly, the true capillaries, which maintain their normal intermittent circulation because of the persistence of the precapillary sphincteric action, continue to perform their function of inward filtration undisturbed by the histamine. These two reactions would explain the lower protein content of the venous blood.

Adrenal cortical extracts. Much indirect evidence has accumulated to indicate that cortical extracts have a counteracting effect on the increase in capillary permeability produced by many noxious agents. Furthermore, in adrenal cortical insufficiency an increase in permeability is indicated by a pathological disturbance of fluid-exchange in favor of outward filtration and the production of a marked hemoconcentration (cf. 43). The question remains as to whether the observed effects are to be explained by the regulatory action of the cortical hormones on the distribution of blood to the tissues or whether these hormones serve to maintain the structural integrity of the capillary wall.

Hechter (cf. 61), by applying histamine through a stab in the skin of rats, observed no difference between normal and adrenalectomized rats in the extravascular loss of trypan blue into the region. Cope and co-workers (cf. 44) obtained an increased protein content in the cervical lymph of adrenalectomized dogs. They claimed this to indicate an increase in permeability of the blood capillaries to protein.

We have found, by direct observation of the capillary circulation, that bilaterally adrenalectomized rats (48 to 96 hrs. post-operative) show no visible alterations in the endothelial wall. Pathological changes of the capillary wall were detected only just prior to circulatory collapse preceding death. It is known that adrenalectomy induces the development of hemoconcentration. This might indicate increased capillary permeability. However, the excessive loss of fluid from the circulation can be explained equally well by an alteration of the hemodynamic relations of the capillary bed. We have found that adrenalectomized rats soon develop an atonic state of the peripheral arteries and arterioles and, subsequently, a cessation of vasomotion in the capillary bed. The resulting derangement of the capillary blood flow is one which favors outward filtration. The failing vasomotion and the consequent inadequate venous return, leading to increasing back pressure in the capillary bed, could account for loss to the tissues not only of water but also of proteins without considering that the integrity of the capillary wall has been impaired.

In the present article no attempt has been made to cover experimental findings dealing with the physiology of the more highly specialized components of the peripheral vascular apparatus, such as those dealing with temperature-regulating

mechanisms in the skin, special secretory barriers such as are obtained in the brain and complex structures observed in the liver and spleen, or in the bat's wing.

Although much has been written on permeability-changes in the capillary bed, most of the findings are largely inferential and only a few are based on direct observations of the vessels. Many of the findings are significant but, until the actual site of their action is ascertained, cannot be used for an understanding of the mechanisms involved.

The fact must be appreciated that the capillary bed has a highly specialized organization the several components of which maintain a delicately adjusted functional balance between reactions of central and local origin. The dominance of a central control tends to be countered by local autonomous control. Along with this ever-shifting balance of the hemodynamics are possibilities of physicochemical changes in the composition of the several different structures which constitute the wall of a capillary. The action of specific agents on one or more of the various components, structural and chemical, requires further intensive study.

Unless adequate precautions are taken and direct observational work is adhered to, it is generally impossible to differentiate between the effect of shifts in hemodynamic relationships, changes in functional permeability of the wall of capillaries, or actual disruptive leakages of the vessels. Most of the data in the literature deal with overall changes involving one or more of the above phenomena. Their physiological significance with respect to the control of capillary permeability will remain obscure until the site of action of the various agencies can be determined.

REFERENCES

(1) Chambers, R. The relation of extraneous coats to the organization and permeability of cellular membranes. C. S. H. Symp. Quant. Biol. 8: 144, 1940.
(2) Keosian, J. Secretion in tissue cultures. III. Tonicity of fluid in chick mesonephric cysts. J. Cell. Comp. Physiol. 12: 23, 1938.
(3) Hoeber, R. Ueber Resorption im Darm. Pflüger's Arch. 86: 199, 1901.
(4) Krogh, A. The active and passive exchanges of inorganic ions through the surface of living cells and membranes in general. Croonian Lecture, Proc. Roy. Soc. Series B, 133: 140, 1946.
(5) Florey, H. Capillary permeability. J. Physiol. (Proc.) 61: 1, 1925.
(6) Field, M. E. and C. K. Drinker. The passage of visible particles through the walls of the capillaries. Am. J. Physiol. 116: 597, 1936.
(7) Abell, R. G. The permeability of blood capillary sprouts and newly formed blood capillaries as compared to that of older capillaries. Am. J. Physiol. 147: 237, 1946.
(8) Wilbrandt, W. Physiologie der Zell- und Kapillarpermeabilität. Helv. Med. Acta 13: 143, 1946.
(9) Danielli, J. F. and A. Stock. The structure and permeability of blood capillaries. Biol. Rev. 19: 81, 1944.
(10) Chambers, R. and B. W. Zweifach. Capillary endothelial cement in relation to permeability. J. Cell. and Comp. Physiol. 15: 255, 1940.
(11) Landis, E. M. Capillary pressure and capillary permeability. Physiol. Rev. 14: 404, 1934.

(12) Brinkmann, R. and A. Szent-Györgyi. Studien über die physikalisch-chemischen Grundlagen der vitalen Permeabilität. I. Die Wirkung kapillaraktiver Stoffe auf die Permeabilität von Kollodium-Membranen. Pflüger's Arch. **245**: 22, 1941.

(13) Danielli, J. F. Capillary permeability and edema in the perfused frog. J. Physiol. **98**: 109, 1940.

(14) Krogh, A. and G. A. Harrop. On the substance responsible for capillary tonus. J. Physiol. **54**: Proc. cxxv, 1921.

(15) Drinker, C. K. The permeability and diameter of the capillaries in the web of the brown frog when perfused with solutions containing pituitary extract and horse serum. J. Physiol. **62**: 249, 1927.

(16) Heimburger, H. Beiträge zur Physiologie der menschlichen Kapillaren. V. Farbenversuche am Kapillarendothel und die Lymphräume des Papillarkörpergewebes. Ztschr. f. d. ges. exper. Med. **55**: 17, 1927.

(17) Benninghoff, A. Ueber die Formenreihe der glatten Muskulatur und die Bedeutung der Rouget'schen Zellen der Kapillaren. Ztschr. f. Zellforsch. u. mikr. Anat. **4**: 125, 1926.

(18) Midsumo, R. Beiträge zur Morphologie und Physiologie der terminalen Blutbahn. Beitr. z. path. Anat. u. allg. Path. **84**: 183, 1930.

(19) Michels, N. A. The structure of the capillaries and the unmyogenic character of Rouget cells (pericytes) in the omentum of rabbits and in the web of living frogs. Anat. Rec. **65**: 99, 1936.

(20a) Volterra, M. Einige neue Befunde über die Struktur der Kapillaren und ihre Beziehungen zur sogenannten Kontraktilität derselben. Zentralbl. f. inn. Med. **46**: 876, 1925; Ricerche sul Sistema redicoloistiocitario. Lo Sperimentale **81**: 319, 1927; Ulteriori studi sulla morfologia e sulla funzione dei capillari sanguigni. Arch. ital. di Anat. e di Embriol. **33**: 844, 1934.

(20b) Volterra, M. Ueber die Pathogenese der Nierenblutungen und ihre anatomischen Ursachen. Zentralbl. f. inn. Med. **36**: 857, 1926.

(21) Gorev, N. N. and A. J. Smirnova-Zamkova. Observations on the pathogenesis of hypertension. Am. Rev. Sov. Med. **3**: 28, 1945.

(22) Duran-Reynals, F. Tissue permeability and spreading factor in infection. Bact. Rev. **6**: 197, 1942.

(23) McClean, D. A factor in culture filtrates of certain pathogenic bacteria which increases permeability of the tissues. J. Path. and Bact. **42**: 477, 1936.

(24) Baron, H. and R. Chambers. A micromanipulative study on the migration of blood cells in frog capillaries. Am. J. Physiol. **114**: 700, 1936.

(25) Chambers, E. L. Unpublished data.

(26) Chambers, R. and B. W. Zweifach. Topography and function of the mesenteric capillary circulation. Am. J. Anat. **75**: 173, 1944.

(27) Landis, E. M. Micro-injection studies of capillary permeability. II. The relation between capillary pressure and the rate at which fluid passes through the walls of single capillaries. Am. J. Physiol. **82**: 217, 1927.

(28) McMaster, P. D. Intermittent take-up of fluid from the cutaneous tissue. J. Exper. Med. **73**: 67, 1941.

(29) Starling, E. H. The fluids of the body. London, A. Constable & Co., 1909.

(30) Landis, E. M. Capillary pressure and hyperemia in muscle and skin of frog. Am. J. Physiol. **98**: 704, 1931.

(31) Landis, E. M. The capillary blood pressure in mammalian mesentery as determined by the micro-injection method. Am. J. Physiol. **93**: 353, 1930.

(32) Landis, E. M. Micro-injection studies of capillary blood pressure in human skin. Heart **15**: 209, 1930.

(33) Landis, E. M. The capillary pressure in frog mesentery as determined by micro-injection methods. Am. J. Physiol. **75**: 548, 1926.

(34) Gordon-Königes, H. and M. Otto. Studies on the filtration mechanism of the intestinal lymph and on the action of acetylcholine on it and on the circulation of the intestinal villi. Quart. J. Exper. Physiol. 26: 319, 1937.

(35) Rous, P., H. P. Gilding and F. Smith. Gradient of vascular permeability. J. Exper. Med. 51: 807, 1930.

(36) Landis, E. M. Capillary permeability and the factors affecting the composition of capillary filtrate. Annals N. Y. Acad. Sc. 46: 713, 1946.

(37a) Smith, F. and M. Dick. Influence of plasma colloids on gradient of permeability. J. Exper. Med. 56: 371, 1932.

(37b) McMaster, P. D., S. Hudack and P. Rous. The relation of hydrostatic pressure to the gradient of capillary permeability. J. Exper Med., 55: 203, 1932.

(38) Zweifach, B. W. The structural basis of permeability and other functions of blood capillaries. C. S. H. Symp. Quant. Biol. 8: 216, 1940.

(39) Keys, A., H. L. Taylor, O. Michelson and A. Henschel. Famine edema and the mechanism of its formation. Science 103: 669, 1946.

(40) Ralli, E. P., J. S. Robson, D. Clarke and C. L. Hoagland. Factors influencing ascites in patients with cirrhosis of the liver. J. Clin. Investigations 24: 316, 1945.

(41) Engel, D. The influence of the sympathetic nervous system on capillary permeability. J. Physiol. 99: 161, 1940.

(42) Gellhorn, E. Das Permeabilitätsproblem. Berlin, Springer, 1929.

(43) Swingle, W. W., W. M. Parkins and A. R. Taylor. A study of the circulatory failure of adrenal insufficiency and analogous shock-like conditions. Am. J. Physiol. 123: 659, 1938.

(44) Cope, O., A. G. Brenizer, Jr. and H. Polderman. Capillary permeability and adrenal cortex studies of cervical lymph in the adrenalectomized dog. Am. J. Physiol. 137: 69, 1942.

(45) Menkin, V. Effect of adrenal cortex extract on capillary permeability. Am. J. Physiol. 129: 691, 1940.

(46) Freed, S. C. and E. Lindner. The effect of steroids of the adrenal cortex and ovary on capillary permeability. Am. J. Physiol. 134: 258, 1941.

(47) Shleser, I. H. and S. C. Freed. The effect of peptone on capillary permeability and its neutralization by adrenal cortical extract. Am. J. Physiol. 137: 426, 1942.

(48) Hyman, C. and R. Chambers. Effect of adrenal cortical compounds on edema formation of frog's hind limbs. Endocrinology 32: 310, 1943.

(49) Graham, J. S. Effect of carbon arc irradiation and adrenal cortical preparations on capillary permeability. Proc. Soc. Exper. Biol. and Med. 54: 101, 1943.

(50) Fine, J. and J. Fischmann. A study of the effect of desoxycorticosterone acetate on capillary permeability. Proc. Soc. Exper. Biol. and Med. 49: 98, 1942.

(51) Swingle, W. W. and J. W. Remington. The effect of DCA and of blood serum transfusions upon the circulation of the adrenalectomized dog. Am. J. Physiol. 134: 503, 1941.

(52) Hechter, O., L. Krohn and J. Harris. Effects of estrogen and other steroids on capillary permeability. Endocrinology 30: 598, 1942.

(53) Rigdon, R. H. and R. B. Chrisman, Jr. Effect of alpha estradiol benzoate on local areas of inflammation in the skin of the rabbit. Endocrinology 28: 758, 1941.

(54) Silver, A., I. E. Stech and C. I. Reed. A study of the effects of vitamin D on capillary permeability by use of the dye T-1824. J. Lab. Clin. Med. 29: 48, 1944.

(55) Rigdon, R. H. Demonstration of a capillary permeability factor in tissue extracts from normal rabbits. Arch. Surg. 41: 96, 1940.

(56) Menkin, V. Studies in inflammation. J. Exper. Med. 64: 485, 1936.

(57) Duran-Reynals, F. A general permeability increasing effect of a factor from mammalian testicles on blood capillaries. J. Biol. Med. 11: 601, 1939.

(58) Rocha e Silva, M. and C. A. Dragstedt. Nature of the capillary permeability factor present in extracts of normal tissues. Proc. Soc. Exper. Biol. and Med. 46: 303, 1941.

(59) McCarrell, J. D. and C. K. Drinker. Cervical lymph production during histamine shock in the dog. Am. J. Physiol. 133: 64, 1941.

(60) Stead, E. A., Jr. and J. V. Warren. The effect of the injection of histamine into the brachial artery on the permeability of the capillaries of the forearm and hand. J. Clin. Investigation 23: 279, 1944.

(61) Hechter, O. Effect of histamine upon capillary permeability in the skin and muscle of normal and adrenalectomized rats. Endocrinology 32: 135, 1943.

(62) Lange, K. and S. E. Krewer. The dermo-fluorometer. J. Lab. Clin. Med. 28: 1746, 1943.

(63) Fine, J. and A. M. Seligman. A study of the problem of the lost plasma in hemorrhage, tourniquet and burn shock by the use of radioactive iodo-plasma protein. J. Clin. Investigation 23: 720, 1944.

(64) Cope, O. and F. D. Moore. A study of capillary permeability in experimental burns and burn shock using radioactive dyes in blood and lymph. J. Clin. Investigation 23: 241, 1944.

(65) Lewis, T. Blood vessels of the human skin and their responses. London, Shaw and Sons, Ltd., 1927.

(66) Lewis, T. and R. T. Grant. Vascular reactions of the skin to injury. II. Liberation of a histamine-like substance in injured skin. Heart 11: 209, 1921.

(67) Clark, E. R. Intercellular substance in relation to tissue growth. Annals N. Y. Acad. Sc. 46: 733, 1946.

SOME PHYSIOLOGICAL EFFECTS OF CURARE AND THEIR APPLICATION TO CLINICAL MEDICINE

A. R. McINTYRE

Department of Physiology and Pharmacology, University of Nebraska, College of Medicine, Omaha

The revival of interest in curare and the important place it has lately won in medicine and surgery is likely to obscure the fact that the study of South American arrow poisons was among the earliest of toxicological investigations. It was largely out of these studies that modern physiology took root and grew. The complex story of the history, botany and chemistry of the curares is remarkable and has interested physicians for centuries. Unfortunately it will not be possible to enter into a discussion of these matters here; they have, however, recently been described in detail elsewhere (1).

It is the purpose of this review to present the gist of to-day's knowledge of the physiology and pharmacology of curare, to outline some of the implications of this knowledge regarding neuromyic transmission and to describe the practical application of this drug to clinical usage. It should first be pointed out, however, that a large variety of South American arrow poisons are indiscriminately referred to as "curares," many of which differ markedly in composition. Comparatively little experimental work has been performed with the pure alkaloids obtainable from them, and of these, only two have been widely used. These are curarine, a potent amorphous alkaloid of unknown chemical structure isolated by Boehm (2), and d-tubocurarine, a bisbenzylisoquinoline compound first obtained in a pure state from tube-curare by King (3), who worked out its chemical structure. Subsequently d-tubocurarine was obtained by Wintersteiner and Dutcher (4), of the Squibb Institute, from *Chondrondendron tomentosum*, a member of the Menisperm or moonseed family, and this alkaloid, accordingly, has been rendered readily available for experimental work. Almost all the clinical work with curare during the last seven years has been performed with Intocostrin, a partially purified and biologically standardized solution of certain varieties of crude curare; the precursor of this preparation was first prepared in the author's laboratory.

The classical effect of curare. In spite of the enormous amount of experimentation performed during the ninety years which have elapsed since Claude Bernard (5) wrote, "Le Curare, qui anéanit l'action nerveuse sur les muscles, conserve au contraire plus longtemps la contractilité musculaire. Preuve que ce sont là deux actes bien distincts," the exact mode of action of curare remains a matter of dispute. The locus of action is generally agreed upon and is thought to reside in the somewhat ill-defined region where nervous tissue ends and muscle tissue begins. The question as to curare's mode of action will, of course, remain a bone of contention until there is complete agreement among physiologists and pharmacologists concerning the *modus operandi* of neuro-muscular transmission. This is not the place to plead the cause of either of the two chief theories of indirect

excitation of muscle, but rather to examine how well the established facts concerning curarization fit the theories. Quite understandably, Claude Bernard's localization of the point of action of curare at the neuromyic junction focussed attention of investigators on this region and it was Bernard's pupil, Kühne (6), who scrutinized nerve endings minutely in an effort to find answers to the yet unanswered questions concerning curare's exact pharmacological action. Out of Kühne's work came a knowledge of the anatomy of the motor nerves and their peculiar endings now known as "motor end-plates." It was found that in the lizard, during life, the motor end-plates were thin, translucent and almost invisible but, after death or after the application of strong curare solutions, the plates became more opaque and clearly visible. Since the days of Kühne a large number of investigators have examined the motor endings in many species. There is much variation in these structures and in some varieties of frogs there are no large terminal plates, the motor nerves "run into the muscle" with a few small knob-like structures scattered along their terminal twigs, thus Langley (7) maintained that in the frog there was no evidence of the presence of any tissue which could not be identified with certainty as either nerve or muscle. Most, but not all, investigators (8) are agreed that high concentrations of curare will cause changes in the appearance of motor nerve terminations. Recently Carey (9) has described many interesting effects on the motor end-plates, not only as the result of the application of strong curare solutions but also following electrocution and poisoning with carbon monoxide, tetraethyl lead, and prostigmine. Some experimenters have reported that, with lower concentrations of curare, morphological changes in the end-plates of curarized muscle are not found, and King and Willard (10) using d-tubocurarine in adequate concentrations to produce curarization, could find no change in the appearance of the motor end-plates; it was not possible for them to differentiate, from their histological appearance, which specimens were curarized and which were not. Accordingly, while it is undoubtedly true that curare and other poisons are capable of affecting visible changes in the end-plates, these changes are not necessarily specific for curarization. Furthermore, these changes have not, as yet, been demonstrated to be reversible.

Langley's conclusion from experiments in the first few years of the present century are hardly reconcilable with the theory that curarization is brought about by an effect on the motor end-plates. It will be recalled that he found that, in denervated muscle, nicotine will cause contractions, and he showed that these nicotine-provoked contractions are preventable by curare. As mentioned above, Langley found that in the frog there is no difficulty in deciding which tissues were nerve and which tissues were muscle. Hence, in a denervated preparation, the effect of curare was evidently on muscle-tissue itself. Langley's extensive experiments with curare and other drugs led eventually to his "receptor" theory of drug action and his work was an important factor in the subsequent acceptance, by many, of the chemical transmitter theory. However, the concepts of Elliot, Dixon, and Loewi were needed before the chemical transmission theory of neuromyic transmission was extended to striated muscle by

Dale and his co-workers. In the meantime the work of the Lapicques emphasized the electrical aspects of excitability. An electrical transmission mechanism is implicit in their theories and their work strongly influenced physiological thought concerning neuro-muscular physiology for many years.

Assuming for the moment that neuromyic transmission is accomplished simply by the shift in electrical potential accompanying the nerve impulse—*curarization could then be due either to a suppression, partial or complete, of nerve* impulses at their terminals, or to a change in the muscle brought about by curare which results in a raised electrical threshold. Obviously there is considerable difficulty in determining whether or not the ultimate terminals of motor-nerve conduct impulses. Kühne's (11) well known attempts to solve this question by the use of the branching motor-nerves in the frog's gracilis muscle which enabled him, by means of a longitudinal cut, to divide the muscle in such a manner that one portion could be soaked in "strongest curare solution" while the other portion remained free of curare, convinced him that the motor nerves, including their endings, were entirely unaffected. Unpublished investigations made in the author's laboratory have shown that d-tubocurarine in concentrations up to a hundred times greater than those necessary for curarization, are powerless to affect nerve-conduction in the frog sciatic, this was true after 18 hours' exposure of the nerve to the poison—but it has not yet been definitely shown whether or not curare in some way prevents the spread of the impulse from the axon over the terminal surface of the nerve. Kuffler (12) has reported experiments that show, following the application of curare to the myo-neural junction, a progressive decrease in the end-plate spike-potential. He points out that unless the leads are placed in close proximity to the end-plate region it is not possible to detect this diminution. As Kuffler himself remarks, "it would be of interest to know the threshold e.p.p. required for a propagated impulse." It is important to observe that the end-plate potential is depressed, not abolished, by curare.

The second possibility, that curare raises the electrical threshold of muscle, brings us at once to the much disputed theory of Lapicque. It is unnecessary to discuss this theory at length, nor is it necessary to present the views of Lapicque's opponents. This dispute has been reviewed in this journal by Davies and Forbes (13). It will suffice to say that neither the doctrine of isochronism for normal nerve and muscle nor the belief that heterochronism exists in curarization is acknowledged universally. This should not be allowed to detract from the usefulness of Lapicque's concept of chronaxie, nor should it be forgotten that Lapicque's theory of curarization (14) has been of great value in stimulating investigations in nerve and muscle physiology during most of the first half of this century. However, in view of the divergent views regarding the strength-duration curves of muscles and their motor nerves, and the many disagreements regarding the validity of the various techniques employed, (for details of which see the review by Davis and Forbes referred to above) it is obvious that any theory of curarization based on changes in *electrical* threshold of muscle is likely to meet with much justifiable criticism.

According to proponents of the chemical theory of neuromyic transmission,

curarization results either from failure of production of transmitter or from failure of the transmitter to accomplish its normal effect upon muscle. There is no evidence at all for the former but considerable evidence for the latter view. This evidence is as follows: Dale and co-workers have shown that muscles perfused with physiological solutions, when stimulated by their motor nerves, either before or during curarization, liberate an acetylcholine-like substance into the perfusate (15). In the writer's laboratory we have succeeded in showing that, in the eserinized dog, the venous blood from muscles, stimulated by their motor nerves, both before and during curarization, contains a substance capable of stimulating rabbit gut *in vitro* and arresting the perfused frog heart. These effects on the rabbit-gut and frog-heart by the venous blood are prevented by atropine. The substance contained in venous blood from stimulated muscles was found to differ in its effects upon the test tissues from those caused by excess potassium; thus in control experiments the effects of eserinized venous muscle blood, to which potassium was added, were compared with the effects obtained with those of eserinized blood obtained from stimulated muscles, the effects of the latter were more readily abolished by atropine than those of the former (16). Further evidence that curarization cannot be explained on the basis of failure in the production of the transmitter is provided from experiments in which acetylcholine is injected directly into an artery. As is well known, muscle-contractions closely simulating those obtained by stimulation of the motor nerve may thus be obtained. Following the injection of curare subsequent injections of acetylcholine are less effective (17, 18, 19), an observation obviously indicative that the action of curare is one of suppression of the effect of the transmitter rather than its failure of production. In denervated muscle the intra-arterial injection of acetylcholine readily arouses muscle-contractions and larger amounts of acetylcholine when so injected may cause muscle contracture accompanied by electrical silence. Curare prevents the action of acetylcholine in denervated muscles, and d-tubocurarine, when injected rapidly in sufficient concentration, is capable of first increasing the fibrillation and may then cause contractures of denervated muscle; during the contractures the spontaneous fibrillations are suppressed or abolished, and during this time muscles will not respond to acetylcholine (20). These findings are supporting evidence for Langley's belief that the site of action of curare is the muscle "effector substance" (see above), and prove beyond question that curare exerts an action on muscles in the absence of their nerves.

Pharmacological antagonists to curare. Eserine and neostigmine and related substances containing a methyl-carbamic ester group abolish the action of curare. These substances have been extensively examined by Briscoe (21) and others (22). They are known to inhibit the activity of serum acetylcholine esterase and this fact is frequently used to explain their pharmacological actions. This explanation is, unfortunately, not entirely adequate as is indicated by the following facts. Wilson and Wright (23) have shown that acetylcholine injected intra-arterially into curarized muscles has but little effect; however, the combined effect of the injection of acetylcholine and prostigmine is greater than either in-

jected alone. These authors have also shown that potassium greatly accelerates the decurarization effect of prostigmine. We find that the potency of Intocostrin (which inhibits acetylcholine esterase) is not measurably inferior to that exhibited by the same quantity of d-tubocurarine contained in it, although d-tubocurarine itself has little or no inhibitory effect on the above enzyme (24).

The antagonism of calcium to curare is of remarkable interest and it has been shown by Eccles and co-workers (25), Feng (26), and others (27), that the restoration of indirect excitability following the local application of calcium to the curarized nerve-muscle junction is accompanied by an increase in the spike potential of the end-plate region. However, it has also been shown that the exposure of a nerve muscle preparation to a solution *low* in calcium increases enormously the sensitivity of the muscle to the intra-arterial injection of acetylcholine (28). On the other hand, it is said that in the intact animal, the administration of parathyroid hormone raises the tolerance for curare (29). Dihydrotachysterol also is alleged to have the same effect (30). On the basis of present knowledge these somewhat paradoxical observations are difficult to harmonize. However, unreported experiments from this laboratory indicate that calcium-lack decreases the resting potential of motor nerve and that this decrease in resting potential is closely associated with the increase in sensitivity of the muscle to acetylcholine, as observed by Kuffler (1944). Furthermore, this increase in sensitivity to acetylcholine is the probable cause of the repetitive discharges following a single nerve volley such as possibly occurs in some forms of tetany. The beneficial effect of curare in tetany may be, in part, due to the raised threshold of muscle to acetylcholine which reduces, or perhaps abolishes, repetitive muscle responses. However, according to Kuffler (1944), curarine does not prevent the twitches of frog muscle obtained *in vitro* upon exposure of the muscle to low calcium; on the contrary we find in preliminary and unreported experiments that the addition of d-tubocurarine to the low calcium solution decreases the spontaneous activity of frog muscle immersed in it.[1]

The antagonistic effect of epinephrine upon curarization has been attributed to the effect of this substance on tissue permeability (31). However, according to Mies (32), the effect is indirect and is in part due to the action of epinephrine on the blood vessels in close proximity to the nerve-muscle junction. Certainly augmentation of blood-flow in muscle is likely to shorten the duration of action of curare. Guanidine, one of the most powerful decurarizing agents (33), appears, like calcium, to raise the spike-potential in the vicinity of the motor end-plate when single stimuli are applied to the nerve.

The reviewer has advanced the hypothesis that many of the effects of curare, acetylcholine, and other substances possessing pharmacological actions at the region of the junction between muscles and their motor nerves (34), can be harmoniously composed by the assumption that the function of the motor nerve is, in reality, inhibitory to the muscle fibres which otherwise undergo spontaneous

[1] Subsequently reported, see Wendt, Bernard F., and A. R. McIntyre. Proc. Am. Physiol. Soc. 6: 224, 1947.

and rhythmic contractions. This inhibition is attributed to the positive (resting) charge on the axon terminal surface which, by its presence, is supposed to favor acetylcholine esterase activity in the region and effectively prevent the accumulation of significant concentrations of acetylcholine. When the surface of the nerve termination becomes more negative as a result of the arrival of a motor impulse, or as a consequence of injury, a reversal of the activity of the enzyme system is now hypothecated, and choline is thus acetylated. The resulting acetylcholine in turn effects the muscle "receptor surface" which is depolarized and the muscle contracts. This hypothesis accounts for the interrelationship between calcium ion concentration and acetylcholine sensitivity, because, according to this concept, a reduction of the resting potential of the motor end-plate would cause an increase in the concentration of acetylcholine, thereby increasing the effectiveness of any acetylcholine applied. We find that increasing the concentration of calcium ions in the fluids surrounding nerve, increases the positive resting potential at its surface. According to the above hypothesis this would tend to keep the activity of acetycholine esterase high and consequently the concentration of acetylcholine at a minimum, and may, if the concentration of calcium be sufficient, result in a block of neuro-muscular transmission. According to this hypothesis the spontaneous fibrillation of denervation should be interrupted by curare, but Solandt and Magladery (35) report that neither anesthetics such as ether and the barbiturates nor atropine, curare and eserine prevent the fibrillation of denervation; we find that d-tubocurarine injected intra-arterially into denervated fibrillating muscles will, when injected sufficiently rapidly, interrupt spontaneous fibrillation as electrically recorded by standard cathode ray techniques. (See ref. 20.) Feinstein et al. (36) have shown that it is very important to control the temperature of muscles in such experiments, and some of the above discrepancies may possibly be accounted for on this basis. In our experiments the blood supply was interrupted only momentarily and significant temperature changes did not occur.

Effects of curare on the autonomic nervous system and central nervous system. Interest in the classical action of curare has resulted in other pharmacological aspects of curare being somewhat neglected, but these are important. For example, curarine applied to autonomic ganglia is capable of blocking synaptic transmission in these structures (37). With concentrations of d-tubocurarine adequate for curarization of striated muscle, in our experience, this effect occurs rather slowly (38). The interruption of transmission at autonomic ganglia appears to be the result, not of interference with the production of acetylcholine, but due to interruption of its effectiveness at the surface upon which it normally acts. This is strongly indicated by experiments performed by us in which a frog heart was perfused with the perfusate from a turtle heart. The concentration of ions in the solution used for perfusion was carefully adjusted so as to be compatible with both the (donor) turtle heart and the (recipient) frog heart, though, unavoidably, less so than solutions best suited to one or the other. When the turtle vagus was stimulated the arrest of the turtle heart was promptly followed by arrest of the (recipient) frog heart. By diverting the perfusate

from the (recipient) frog heart during curarization of the donor heart and then, during the stimulation of the turtle vagus, again allowing the perfusate to flow through the frog heart, it was sometimes possible to obtain arrest of the recipient heart without arrest of the donor heart. This strongly indicates that although stimulation of the vagus produced *vagus-stoffe* in adequate quantities to arrest the recipient heart it was unable to arrest the curarized turtle donor heart. It should be pointed out that generally both the donor and recipient hearts were arrested upon stimulation of the turtle vagus. Arrest of the recipient heart *only* was seen in experiments in which the application of curare had been prolonged. These and other experiments with curare lead us to believe that, in general, cholinergic structures are more easily affected than adrenergic structures. At first possibly both are stimulated transiently, and later paralyzed, so that autonomic transmission from pre- to post-ganglionic neurones gradually fails and may cease entirely after exposure for sufficiently long periods of time to adequate concentrations of the drug. The effects of curare on the secretions are many. The literature is full of contradictions and inconsistencies. This is because the effects of curare upon the glands are probably mediated by its effects upon the autonomic system and the results obtained depend upon the duration of the experiments and the amount of curare used. Another factor responsible for divergent reports is, no doubt, the differences in the varieties of curare used. The following facts seem to be fairly well established. As a result of curarization, the blood sugar rises. The urine output falls without a significant change in blood-pressure. We suggest, without complete evidence, that the diminished urine secretion is the result of an increase in output of posterior pituitary hormone—the urine during curarization frequently contains much sodium chloride. The tolerance to insulin is increased. Somewhat surprisingly, when the convulsions in febrile conditions are controlled by curarization, the body temperature remains elevated. A large number of observations concerning miscellaneous actions of curare reported by numerous investigators have been collected by the writer (39); space does not permit their quotation here.

The effects of curare on abnormal muscle-tone. The effects of curare on the hypertonicity of muscle in decerebrate animals was first investigated by Bremer (40), who showed that curare was capable of reducing the muscle-tonus in such animals. This was accomplished without the production of classical curarization, the animals being able to walk without noticeable impairment of their gait. Somewhat parallel observations were reported by Hartridge and West (41) in tetany in dogs. Such animals could be protected from tetanic convulsions but at the same time retain their ability to walk, apparently, normally. West (see ref. 31), reported that there was considerable variation in the effectiveness of tetany-control by different curares, some possessing what he termed a "lissive" action, whereas others, including d-tubocurarine, possessed little or no lissive action. This alkaloid, as above mentioned, is the active alkaloid contained in Intocostrin. Consequently it is somewhat surprising to find that Intocostrin and d-tubocurarine are today the preparations chiefly used clinically in the treatment of spasticities. Burman (42), who was the first to report on the

use of curariform drugs in the treatment of spastic patients in this country, made his initial investigations before Intocostrin was available, and the botanical source of the curare he used is not entirely clear. Later Harvey and Masland (43) reported unfavorably on the results of curare in spastic conditions. The writer has had the opportunity of seeing a number of spastic patients who appeared to be benefited by the use of Intocostrin, and others, notably some of Doctor Schlesinger's patients, who seemed undoubtedly improved by the intramuscular injection of d-tubocurarine suspended in oil and beeswax, the effects of which are said to persist for as long as three days (44).

These clinical observations confront the pharmacologist with a difficult problem because in normal dogs the classical peripheral effects of these preparations do not, in the reviewer's experience, persist sufficiently long to account for the duration of the clinical improvements reported. There is also difficulty in explaining Bremer's observation (l.c.) on decerebrate rigidity and those of Hartridge and West (l.c.) in parathyroid tetany in terms of the classical curare effect, because, as mentioned above, these experimenters did not obtain measurable interference with the animals' ability to walk. It would seem that in the event of the curare preventing a certain proportion of the motor-nerve impulses responsible for the increased muscle tonus from reaching the muscles there would be an impairment of muscle-function roughly proportional to the decrease in the hypertonicity. Should it be demonstrated that in spastic states, repetitive muscle responses follow single motor nerve impulses, which seems rather unlikely, the beneficial effects of curare in spasticities would be more readily understandable. During curarization, as is well known, a muscle will rapidly lose its ability to maintain a tetanic contraction and then respond by a single twitch to an indirect tetanic stimulus. This phenomenon, however, is hardly adequate to explain the relief believed to be afforded by curare in hypertonicity. The question is further complicated by the fact that West (l.c.) was unable to find a lissive action with d-tubocurarine. Admittedly knowledge of the numerous factors responsible for pathological hypertonicity is incomplete and accordingly a discussion of the mode of action of curare in these conditions is consequently speculative rather than factual. It seems possible that the effects of curare in spastic states may depend, at least in part, upon some mechanism other than its classical peripheral action. Accordingly attention is drawn to the observations of Denhoff and Bradley (45), who state that their spastic patients experienced subjective manifestations following injection with effective doses of curare. These subjective symptoms might be interpreted as suggestive evidence of a central nervous system action of the drug, and West (46) has quoted the words of one of his patients whose subjective reactions, while under curare therapy, clearly indicated marked disturbances of the sensorium. The pharmacological actions of curare on the brain and cord have been largely neglected but convulsions, following intra-aortic injections, have long been known to occur (47), and when curare gains access to the ventricles marked disturbances of the C.N.S. result (48). Both Harlow (49) and Girden (50) have investigated the effects of curare on conditioned reflexes in dogs and both report changes in the responses of cur-

arized animals which they attribute to central actions. Girden believes that the curarized dog suffers amnesia. The electrical activity of the frog's C.N.S. has been investigated by Pick and his associates (51), who find a depression of activity which may last several days in the curarized animal. We find that the immediate effect of rapid injections of d-tubocurarine into anesthetized dogs is a transient increase in amplitude of the electroencephalographic record promptly followed by a decreased electrical activity of the brain (52). It is possible to cause sudden and complete arrest of respiration and death by the rapid intracarotid injection of d-tubocurarine; respiration may cease before indirect excitability of the diaphragm is lost. It is desirable that the mechanisms responsible for the decrease in muscle hypertonicity observed clinically be investigated with the possible rôle of these actions of curare on the C.N.S. in mind. It is conceivable that the effect obtained in spastic patients involves actions of curare on reflex pathways, resulting in impairment of synaptic transmission, which in the abnormal C.N.S. may be sufficiently modified by curare to account, in part, for the clinical benefits claimed. The subject obviously requires further investigation.

The use of curare in shock-therapy and anesthesia. At the suggestion of the author, A. E. Bennett (53) began the use of a partially purified and carefully standardized solution of curare, prepared in this laboratory, on psychopathic patients undergoing metrazol shock-therapy. The incidence of trauma resulting from the convulsions was reduced to the vanishing point, without diminishing the ratio of successful to unsuccessful treatments. Subsequently with the development of electro-shock therapy the use of curare in shock-therapy has continued and is today used in perhaps a majority of institutions where this type of therapy is employed. It is probable that the dosage of curare for use with electro-shock therapy should be somewhat less than that formerly used in metrazol shock-therapy. The reason for this is that metrazol is antidotal to curare whereas electro-shock-therapy is not. Electroencephalographic records, obtained by us, indicate that both curare and electro-shock result in decreased electroactivity of brain, hence, when both are employed together, they may augment each other.

Griffith and Johnson (54) introduced the use of curare as an adjuvant in anesthesia and Griffith (55) has extensively investigated the usefulness of the drug in this field. Cullen (56) has also investigated curare in anesthesia and Gross and Cullen (57) have made the important observation that with ether the amount of curare necessary to obtain good muscle-relaxation is considerably less than with such anesthetics as cyclopropane. In the opinion of the author cyclopropane anesthesia with curare approaches the ideal in anesthesia for a wide variety of surgical procedures. The deep relaxation obtained with curare supplies the one deficiency of cyclopropane. In abdominal surgery another advantage of curare is that the gut is rendered "small" and inactive. Presumably this effect on the gut is caused by an effect upon synaptic conduction in the autonomic system because we find that the concentrations necessary to modify rabbit jejunum *in vitro* far exceed those likely to be present when the drug is used in man. While the chief value of curare in anesthesia undoubtedly resides in the muscle relaxa-

tion obtained due to the drug's classical peripheral action, there is disagreement as to what rôle, if any, is played by its central actions. There have been a number of papers which report the administration of very large doses of curare to patients, either by accident or by design, in which it appeared that anesthesia resulted, and in some instances major surgery has been successfully accomplished under curare alone (58). On the other hand, Dr. Scott M. Smith (59) allowed himself to be injected with d-tubocurarine to the point of complete curarization without the loss of consciousness. It is possible that the rate of administration of d-tubocurarine in Smith's experiment was insufficiently great to produce the effects reported by others. This important question requires further investigations. It is, however, not too easy to find competent observers who exhibit the devotion to science shown by Doctor Smith.

The use of curare for the control of convulsions in tetanus dates back to the early part of the 19th century and the literature on this topic has been previously reviewed by the author (60). There is no evidence at present that curare is capable of saving the patient's life although control of convulsions can be rather readily accomplished. One of the difficulties in the prolonged use of curare for the maintenance of deep curarization is that, in experimental animals at least, the animal gradually succumbs. The mechanism of death is obscure (61). The reviewer believes that death is related to the gradual interruption of function of the C.N.S. brought about by the prolonged action of curare. In spite of this the experimental use of curare in severe human tetanus is fully justified when adequate facilities for maintaining effective artificial respiration are available. Reliance on atmospheric pressure to fill the lungs, as for example, when an "iron lung" is used, may result in inadequate pulmonary ventilation due in part to bronchial spasm. Methods employing positive pressure should therefore be used.

Ransohoff (62) has taken advantage of curare to decrease the painful muscle spasm frequently found early in poliomyelitis, thereby rendering his patients more comfortable and facilitating the early application of physio-therapy. On the basis of present knowledge it should not be assumed that there is any *direct* beneficial effect exerted by curare.

The only known contraindications to the use of curare are during the acute paralytic stage of poliomyelitis and in *myasthenia gravis*, a disease in which sensitivity to curare is markedly increased. The many parallelisms between partial curarization and *myasthenia gravis* have been described by Bennett, McIntyre and Cash (63). The antidotal value of eserine and prostigmine bromide have been compared by the author (64) and the former was found to be superior; however, subsequent work has shown that prostigmine *methyl sulfate* is fully as effective as eserine.

It is unnecessary to enumerate all the clinical applications of curare which have been attempted; they include the treatment of such diverse conditions as persistent singultus, dysmenorrhea, epilepsy and as an aid in endoscopy. They have been discussed in some detail elsewhere (65). If the claim of the successful treatment of dysmenorrhea is substantiated it is obvious that the practical

pharmacological use of curare is not necessarily limited to its classical effect on skeletal muscle.

SUMMARY

1. The mechanism of curarization cannot be completely explained until the problem of neuro-muscular transmission is fully solved.

2. Curare possesses, in addition to its classical action on neuro-muscular transmission, an action on autonomic synaptic transmission which transmission it is capable of blocking.

3. The effects of a sufficient concentration of a curare on the central nervous system are a transient increase in activity, followed by decreased activity.

4. The above pharmacological actions of curare can be tentatively attributed to its effect upon cholinergic structures, whose ability to respond to acetylcholine is selectively and progressively modified.

5. The clinical use of curare, in shock-therapy and anesthesia, is thoroughly established. The full evaluation of the use of curare in spasticities awaits further experimentation.

6. The clinical use of curare for the treatment of conditions other than those involving the striated musculature is provocative of interest but must be viewed with cautious reserve.

REFERENCES

(1) McIntyre, A. R. Curare, its history, nature, and clinical use. 1947, chapters 1–5.

(2) Boehm, R. Chemische Studien über das Curare. Beitr. g. Physiol., Leipzig, 1886, 173–192.

(3) King, H. Curare alkaloids. I. Tubocurarine. J. Chem. Soc. 1381, 1935.

(4) Wintersteiner, O. and J. D. Dutcher. Curare alkaloids from *Chondodendron tomentosum*. Science 97: 467, 1943.

(5) Bernard, C. Leçons sur les effets des substances toxiques et médicamenteuses. Paris, 1857, 316.

(6) Kühne, W. Ueber die Wirkung des amerikanischen Pfeilgiftes. Arch. f. Anat., Physiol. u. wissensch. Med., 477, 1860.

Kuhne, W. Zur Histologie der motorischen Nervenendigung, Untersuch. a. d. physiol. Inst. Heidelberg 2: 187, 1882.

Kühne, W. Über die Wirkung des Pfeilgiftes auf den Nervenstämme. Heidelberg, 1886.

(7) Langley, J. N. On the contraction of muscle chiefly in relation to the presence of receptive substances. J. Physiol. 36: 347, 1907. Ibid. 37: 165, 1908. Ibid. 38: 285, 1908. Ibid. 39: 235, 1909.

(8) Judée. Mode d'action du curare sur les cellules nerveuses dites motrices ou grosses cellules. Compt. rend. Soc. d. biol., 8.s., 1: 292, 1884.

(9) Carey, E. J. Effect of CO_2 on ameboid changes in motor nerve plates in intercostal muscle. Proc. Soc. Exper. Biol. and Med. 47: 67, 1941.

Carey, E. J. Pathologic effects of CO_2 and electricity on the explosive ameboid motion in motor nerve plates in intercostal muscle. Am. J. Path. 18: 237, 1942.

Carey, E. J. Experimental histopathology of motor end plates produced by quinine, curare, prostigmine, acetylcholine, strychnine, lead tetraethyl, and heat. Ibid. 20: 341, 1944.

(10) King, R. and W. A. Willard. The effects of fatigue and curare on the morphology of the motor end plates. Anat. Rec. 91: 286, 1945 (abstract).

(11) KÜHNE, W. Utersuchungen über Bewegungen und Veränderungen der contraktilen Substanzen. Arch. f. Anat., Physiol. u. wissensch. Med. 564–642, 1859.

(12) KUFFLER, S. W. Electric potential changes at an isolated nerve-muscle junction. J. Neurophysiol. 5: 18, 1942.

(13) DAVIS, H. AND A. FORBES. Chronaxie. Physiol. Rev. 16: 407, 1936.

(14) LAPICQUE, L. Neuromuscular isochronism and chronological theory of curarization. J. Physiol. 81: 113, 1934.

(15) DALE, H. H., W. FELDBERG AND M. VOGT. Release of acetylcholine at voluntary motor nerve endings. J. Physiol. 86: 353, 1936.

(16) McINTYRE, A. R. Curare, its history, nature, and clinical use. Univ. of Chicago Press, 1947, chapter 8.

(17) DEHENNOT, O. Les poisons curarisants et la fatigue musculaire. Compt. rend. Soc. de biol. 117: 318, 1934.

(18) SIMONART, A. AND F. E. SIMONART. L'acétylcholine et le muscle strié normal de mammifère, Arch. internat. de pharmacodyn, et de thérap. 49: 302, 1935.

(19) SIMONART, A. Etude de l'action de l'acétylcholine sur le muscle strié de mammifère. Arch. internat. de pharmacodyn, et de thérap. 51: 381, 1935.

(20) McINTYRE, A. R., R. E. KING AND A. L. DUNN. Electrical activity of denervated mammalian skeletal muscle as influenced by d-tubocurarine. J. Neurophysiol 8: 297, 1945.

(21) BRISCOE, G. Antagonism between curarine and prostigmin and its relation to myasthenia gravis. Lancet 1: 470, 1936.

BRISCOE, G. Depressant peripheral effect of prostigmin and its temporary antagonism by curarine. J. Physiol. 86: 1P, 1936.

BRISCOE, G. The antagonism between curarine and acetylcholine. Ibid. 87: 425, 1936.

BRISCOE, G. Anti-curare action of substance 36, closely related to prostigmin. Lancet 1: 621, 1937.

BRISCOE, G. Eseriniform and curariform action of certain onium salts. J. Physiol. 93: 6P, 1938.

BRISCOE, G. Changes in muscle contraction curves produced by drugs of the eserine and curarine groups. Ibid. 194–205.

BRISCOE, G. The antagonism between curarine and prostigmin. Lancet 1: 469, 1936.

(22) ECCLES, J. C., B. KATZ AND S. W. KUFFLER. Effect of eserine on neuromuscular transmission. J. Neurophysiol. 5: 211, 1942.

(23) WILSON, A. T. AND S. WRIGHT. Anti-curare action of potassium and other substances. Quart. J. Exper. Physiol. 26: 127, 1937.

(24) McINTYRE, A. R. AND R. E. KING. d-Tubocurarine chloride and choline esterase. Science 97: 69, 1943.

(25) ECCLES, J. C., B. KATZ AND S. W. KUFFLER. Electric potential changes accompanying neuro-muscular transmission. Biol. Symposia 3: 349, 1941.

(26) FENG, T. P. Studies on the neuro-muscular junction: universal antagonism between calcium and curarising agencies. Chinese J. Physiol. 10: 513, 1936.

(27) CHAO, J. Action of calcium and curare on muscular contraction and neuromuscular transmission. Chinese J. Physiol. 10: 545, 1936.

(28) KUFFLER, S. W. Specific excitability of the endplate region in normal and denervated muscle. J. Neurophysiol. 6: 99, 1943.

KUFFLER, S. W. The effect of calcium on the neuromuscular junction. J. Neurophysiol. 7: 17, 1944.

(29) KELBLING, W. Protective action of parathyroid hormone against poisoning by curarine. Arch. f. exper. Path. u. Pharmakol. 195: 567, 1940.

(30) WIRTH, F. Protective action of dihydrotachysterol against curarine poisoning. Arch. f. Exper. Path. u. Pharmakol. 195: 564, 1940.

(31) WEST, R. A. A pharmacological study of derivatives of two specimens of *tubo*-curare,

and an examination of four members of genus *Strychnos*. Arch. internat. de pharmacodyn. et de thérap. 56: 81, 1937.

(32) MIES, H. Curare und vegetative Innervation des Skelettmuskels. Ztschr. f. Biol. 98: 70, 1937.

(33) FENG, T. P. Studies on neuro-muscular junction: the effect of guanidine. Chinese J. Physiol. 13: 119, 1938.

(34) McINTYRE, A. R. Curare, its history, nature, and clinical use. 1947, chapter 15.

(35) SOLANDT, D. Y. AND J. W. MAGLADERY. The relation of atrophy to fibrillation in denervated muscle. Brain 63: 255, 1940.

(36) FEINSTEIN, B., R. E PATTLE AND G. WEDDELL. Metabolic factors affecting fibrillation in denervated muscle. J. Neurol. Neurosurg. and Psychiat. 8: 1, 1945.

(37) BROWN, G. L. AND W. FELDBERG. Differential paralysis of the superior cervical ganglion. J. Physiol. 86: 10P, 1936.

(38) McINTYRE, A. R. Curare, its history, nature, and clinical use. 1947, chapter 9.

(39) McINTYRE, A. R. Curare, its history, nature, and clinical use. 1947, chapters 9, 10, 11, 12, 13 and 14.

(40) BREMER, F., J. TITECA AND L. VAN. MEIREN. Sur la sensibilité au curare de la rigidité tétanique locale. Compt. rend. Soc. de biol. 97: 895, 1927.

BREMER, F. AND J. TITECA. Nouvelles recherches sur le mécanisme de l'atone curarique. Compt. rend. Soc. de biol. 107: 253, 1931.

BREMER, F. AND J. TITECA. Mécanisme de l'action électif du curare sur le tonus postural et les hypertonies. Ibid. 97: 1407, 1927.

BREMER, F. AND J. TITECA. Action de l'adrénaline sur l'atonie musculaire du stade initial de la curarisation. Ibid. 99: 624, 1928.

BREMER, F. AND J. TITECA. Atonie curarique et inhibition de Wedensky. Arch. internat. de physiol. 42: 223, 1935.

(41) HARTRIDGE, H. AND R. WEST. Parathyroid tetany in dogs and its abolition by curare. Brain 54: 312, 1931.

(42) BURMAN, M. S. Therapeutic use of curare and erythroidine hydrochloride for spastic and dystonic states. Arch. Neurol. and Psychiat. 41: 307, 1939.

BURMAN, M. S. Curare therapy for the release of muscle spasm and rigidity in spastic paraylsis and dystonia musculorum deformans. J. Bone and Joint Surg. 20: 754, 1938.

BURMAN, M. S. Clinical experiences with some curare preparations and curare substitutes. J. Pharmacol. and Exper. Therap. 69: 143, 1940.

BURMAN, M. S. Curare therapy. J. A. M. A. 129: 1286, 1945.

(43) HARVEY, A. M. AND R. L. MASLAND. Actions of curarizing preparations in the human. J. Pharmacol. and Exper. Therap. 73: 304, 1941.

(44) SCHLESINGER, E. B. Use of curare in oil in treatment of spasticity following injury of the spinal cord. Arch. Neurol. and Psychiat. 55: 530, 1946.

(45) DENHOFF, E. AND C. BRADLEY. Curare treatment of spastic children. New England J. Med. 226: 411, 1942.

(46) WEST, R. The pharmacology and therapeutics of curare and its constituents. Proc. Roy. Soc. Med. 28: 565, 1935.

(47) JOSEPH, D. R. AND S. J. MELTZER. On the difference between the effects of intravenous and intra-aortic injections of curarin in frogs. J. Pharmacol. and Exper. Therap. 3: 465, 1911–12.

(48) STERN, L. AND E. ROTHLIN. Effects of the local application of curare upon different parts of the cerebellum. Arch. Sc. Phys. Nat., Suppl. 11–19, 4: 45, 1918.

STERN, L. AND E. ROTHLIN. Du effets de l'application directe du curare sur les différentes parties du cervelet. Compt. rend. Soc. de phys. et d'hist. nat. 35: 11, 1918.

STERN, L. AND E. ROTHLIN. Effets de l'application directe du curare sur les différentes parties du cervelet. Schweiz. Arch. f. Neurol. u. Psychiat. 3: 234, 1918.

(49) Harlow, H. F. and R. Stagner. The effect of complete striate muscle paralysis on the learning process. J. Exper. Psychol. 16: 283, 1933.
Harlow, H. F. The effects of incomplete curare paralysis upon formation and elicitation of conditioned responses in cats. J. Genet. Psychol. 56: 273, 1940.
(50) Girden, E. and E. Culler. Conditioned response in curarized striate muscle in dogs. J. Comp. Psychol. 23: 261, 1937.
Girden, E. Cerebral mechanisms in conditioning under curare. Am. J. Psychol. 53: 397, 1940.
Girden, E. Generalized conditioned responses under curare and erythroidine. J. Exper. Psychol. 31: 105, 1942.
Girden, E. Dissociation of pupillary conditioned reflexes under erythroidine and curare. Ibid. 322–32.
(51) Feitelberg, S. and E. P. Pick. Action of curare on the brain of the frog. Proc. Soc. Exper. Biol. and Med. 49: 654, 1942.
Pick, E. P. and K. Unna. The effect of curare and curare-like substances on the central nervous system. J. Pharmacol. and Exper. Therap. 83: 59, 1945.
(52) McIntyre, A. R., A. L. Dunn and P. E. Tullar. The effect of d-tubocurarine on the electrical activity of dogs' brains. Federation Proc. 5: No. 1, 67, 1946.
(53) Bennett, A. E., A. R. McIntyre and A. L. Bennett. Pharmacologic and clinical investigations with crude curare, J. A. M. A. 114: 1791, 1940.
(54) Griffith, H. R. and G. E. Johnson. The use of curare in general anesthesia. Anesthesiology 3: 418, 1942.
(55) Griffith, H. R. The use of curare in anaesthesia and for other clinical purposes. Canad. M. A. J. 50: 144, 1944.
Griffith, H. R. Curare in anesthesia. J. A. M. A. 127: 642, 1945. Also Lancet 2: 74, 1945.
Griffith, H. R. Curare in anesthesia (reply). J. A. M. A. 127: 644, 1945.
Griffith, H. R. Curare—a new tool for the anaesthetist. Canad. M. A. J. 52: 391, 1945.
(56) Cullen, S. C. The use of curare for improvement of abdominal muscle relaxation during inhalation anesthesia. Surgery 14: 261, 1943.
Cullen, S. C. Clinical and laboratory observations on the use of curare during inhalation anesthesia. Anesthesiology 5: 166, 1944.
Cullen, S. C. The use of curare in anesthesia. South. M. J. 38: 144, 1945.
(57) Gross, E. G. and S. C. Cullen. The effects of anesthetic agents on muscular contraction. J. Pharmacol. and Exper. Therap. 78: 358, 1943.
(58) Whitacre, R. J. and A. J. Fisher. Use of curare in anesthesia. Ohio State M. J. 40: 1155, 1944. Also: Anesthesiology 6: 124, 1945.
(59) Smith, S. M. Symposium on curare. A. M. A. session, San Francisco, J. A. M. A. 131: 1148, 1946.
(60) McIntyre, A. R. Curare, its history, nature, and clinical use. 1947, chapter 14.
(61) Perlstein, M. A. and A. Weinglass. Fatal effects of prolonged complete curarization. Am. J. Dis. Child. 67: 360, 1944.
(62) Ransohoff, N. S. Curare in the acute stage of poliomyelitis: preliminary report. J. A. M. A. 129: 129, 1945.
(63) Bennett, A. E., A. R. McIntyre and P. T. Cash. Comparison of curarization with myasthenia gravis. Exhibit, A. M. A., J. A. M. A. 131: 321, 1946.
(64) McIntyre, A. R., P. Tullar and M. Frank. Failure of neostigmine as an antidote for curare. Proc. of the Cent. Soc. for Clinical Research 19: 73, 1946.
(65) McIntyre, A. R. Curare, its history, nature, and clinical use. 1947, chapter 14.

THE URINARY COPROPORPHYRINS IN HEALTH AND DISEASE[1]

CECIL JAMES WATSON AND EVREL A. LARSON

Department of Medicine, University of Minnesota Hospital, Minneapolis

The literature relating to the urinary porphyrins has become so extensive that it would scarcely be feasible to review it comprehensively in the space which has been allotted to the present paper. The authors have striven rather to bring together the essential pieces of information which comprise the existing knowledge of the subject. A number of reviews have appeared to which the reader is referred for additional orientation to the literature, both with respect to the urinary porphyrins (9d; 16i; 99c), and to porphyrin metabolism in general (9c; 11; 16i; 24e, k; 32a, b, d; 58b; 73b; 78b; 97c; 100g, h). Because of space restriction, the present review will concern itself mainly with the urinary coproporphyrins. Other porphyrins and other aspects of the coproporphyrin problem will be considered only as they relate to the coproporphyrins of the urine.

The term uro- and coproporphyrin have proven somewhat unfortunate in the light of knowledge of these substances gained since Hans Fischer first described and named them. Prior to Fischer's studies, the naturally occurring porphyrins had been generally assumed to be the hematoporphyrin as first prepared by Hoppe-Seyler (41) and later, in crystalline form, by Nencki and Sieber (64). Reasons for this confusion will appear presently. Fischer's first observations were made in a case of congenital porphyria, the famous case Petry, in which the urine, as it characteristically does in this disease, contained large amounts of uro- and relatively small amounts of coproporphyrin (24a). The feces from this case was found to contain only coproporphyrin (24b), and in fact, the presence of uroporphyrin in the feces of porphyria cases was overlooked until relatively recently (80a). Fischer's isolation of crystalline uro- and coproporphyrin from urine and feces, respectively, initiated a brilliant series of studies of the porphyrins which were terminated by his untimely death in 1945. These studies of Fischer and his school have provided much of the existing knowledge concerning the structure and chemical interrelationships of the porphyrins. Relatively early in the course of this long research, Fischer was able to show that the configuration of the uro- and coproporphyrin isolated from the case Petry, differed from that of the protoporphyrin in the hemoglobin molecule (24e). This difference may be noted in figure 1, in which it is seen that there is opposite placement of the methyl and propionic acid groups on pyrrol nucleus IV (left lower). Fischer synthesized the four "aetioporphyrins", which have only methyl and ethyl groups on the porphyrin ring (24e, k). These do not occur in nature, but they serve as the basis for his classification, in that all other porphyrins belong to one of four main types, depending upon correspondence of the side groups with the methyl or ethyl groups of the aetioporphyrins. Thus the uro- and copro-

[1] Aided by grants from the John and Mary R. Markle Foundation, New York City, and the Medical Research Fund of the Graduate School, University of Minnesota.

porphyrin isolated from the porphyria case were found to have the configuration of aetioporphyrin I, while the protoporphyrin of the hemoglobin molecule was later shown to have the type III structure[2] (24p). Fischer thus revealed the "dualism" of the porphyrins in nature. Although theories of porphyrin synthesis in the animal organism have been proposed which postulate a minute formation of type II or IV porphyrins (16g, i; 94), they have never been isolated from natural material nor has any evidence been described to indicate their occurrence.

As will be discussed subsequently, the formation and excretion of coproporphyrins I and III, varying in relative and absolute amounts under different circumstances, has been well established. Their fundamental significance, however, is but poorly understood. The formation and excretion of uroporphyrin I has come to be recognized as a characteristic feature of the metabolic error

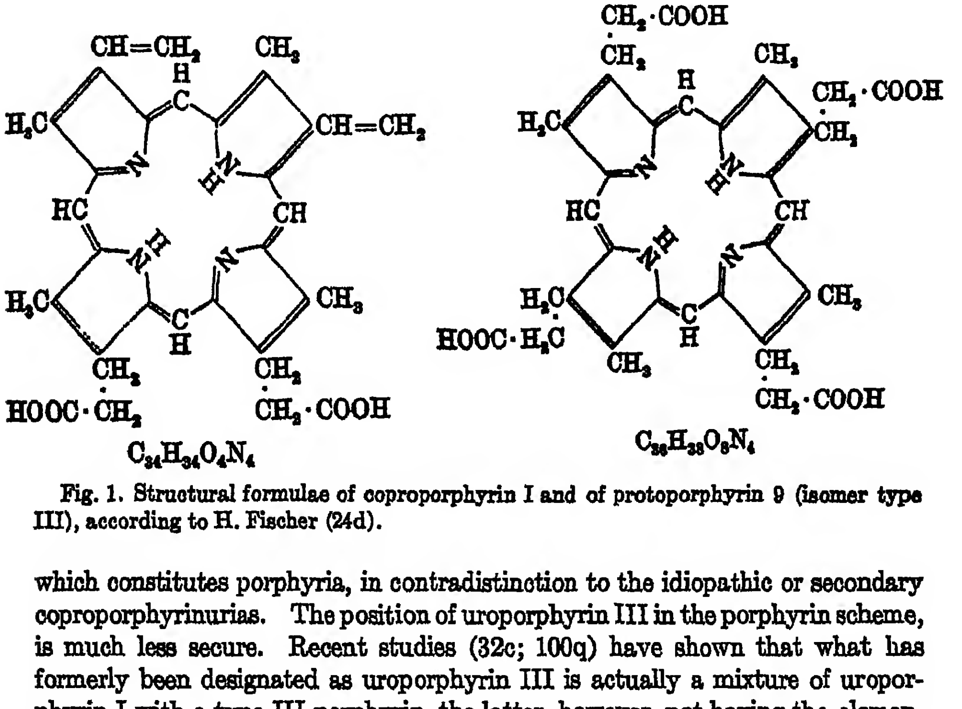


Fig. 1. Structural formulae of coproporphyrin I and of protoporphyrin 9 (isomer type III), according to H. Fischer (24d).

which constitutes porphyria, in contradistinction to the idiopathic or secondary coproporphyrinurias. The position of uroporphyrin III in the porphyrin scheme, is much less secure. Recent studies (32c; 100q) have shown that what has formerly been designated as uroporphyrin III is actually a mixture of uroporphyrin I with a type III porphyrin, the latter, however, not having the elementary composition of a uroporphyrin isomer. The uroporphyrin problem would require separate consideration from many standpoints, and except where it relates in an intimate way to the urinary coproporphyrin, it will not be included in the ensuing review.

[2] Actually this is designated as protoporphyrin 9, since it corresponds in configuration with mesoporphyrin 9. Fischer further classified porphyrins having more than two types of side groups, on the basis of correspondence in configuration with one of the 15 mesoporphyrins. The latter have four methyl, two ethyl, and two propionic acid side groups.

The physical constants of the coproporphyrins and their metal complexes and esters, including melting points, crystal habitus, absorption and fluorescence spectra, and solubilities have been given in detail elsewhere (16i; 24k; 49; 50; 100h).

The coproporphyrin of normal urine. MacMunn (56), Saillet (76) and Garrod (27a, b, c) were the first to recognize the presence of a porphyrin in normal urine. These earlier investigators labored under the difficulty of rather crude spectroscopes with which it was scarcely possible to distinguish between copro- and hematoporphyrin, the absorption spectra of which differ by not more than 10 Å (24k). Thus it is not surprising that the porphyrin of normal urine, as well as that found in increased amounts in a variety of pathological states, was believed for many years to be the hematoporphyrin which Nencki and his co-workers first established as a well defined chemical entity (64). Schumm (79a, b) recognized the differing absorption spectra of the urinary porphyrin and hematoporphyrin, indicating his belief that the former was identical with Fischer's coproporphyrin (24a, b). This spectroscopic identity was confirmed by Fischer and Zerweck (24m). Hoerburger (39a) first purified the coproporphyrin of normal human urine, stating that it was the type I isomer on the basis of the crystal habitus of its methyl ester. Fink and Hoerburger (23) reported that the pH fluorescence curve of this crystalline material was that of coproporphyrin I. Crystals obtained from the mother liquor exhibited the habitus of the type III isomer. Watson (100e) isolated 0.13 mgm. of crystalline ester from 33 liters of normal human urine (one individual). The crystal habitus was that of type I, but the melting point was somewhat low at 228°C. (Coproporphyrin I methyl ester melts at 250°C., while the methyl ester of type III has a dimorphic melting point, i.e., 135°, 144°, 167–170°C. [24i, k].) This together with the fact that a few rosettes of prisms were obtained from the mother liquor suggested the presence of a smaller amount of type III isomer. By means of preliminary concentration on infusorial earth, Grotepass (33b) isolated the coproporphyrin from 10,000 liters of presumably normal human urine, pooled from a large number of individuals. Upon fractional crystallization of the yield, which was approximately 200 mgm., 96 mgm. of type I and 87 mgm. of type III were obtained. As pointed out by Dobriner and Rhoads (16i), the chief objection to Grotepass' conclusion that this is the isomer distribution of normal human urine, is the possibility of inclusion in the pool of urine from an individual having a latent or symptomless type III porphyrinuria. It will be noted subsequently that such a condition may easily be overlooked. Furthermore, it was possible that chemical or metal exposure, or alcoholism (vide infra) was present in at least some of the individuals contributing to Grotepass' pool. It was desirable, therefore, to know the isomer distribution in a series of urines for each of which it had first been determined that the concentration of total coproporphyrin was within the normal range (vide infra). Data of this type have recently been gained (100s) in a study of 30 normal individuals, employing the differential precipitation "fluorescence quenching" technique (80b, e). The range of isomer distribution in this study was from 65 to 92 per cent type I, and from 8 to 35 per cent type III. These 30 individuals

represented both sexes about equally, but were all below 30 years of age, the majority being students between 20–25. An additional group of 11 individuals of both sexes between the ages of 69–82 were also studied (100s). This revealed a slightly higher relative percentage of type III isomer. It appears, however, that this apparent increase is due to a mild decrease of type I. The range of values for the type I and type III isomers in the young and old individuals was as follows:

	YOUNG	OLD
I	16.—89 γ per day	14.7—57 γ per day
III	1.4—31.7 γ per day	4.5—34.3 γ per day

The concentration and per diem output of the total coproporphyrin in normal human urine has been determined repeatedly in the past, the reported values

TABLE 1

The average normal excretion of coproporphyrin in the urine as determined by various investigators

AUTHOR	URINARY COPROPORPHYRIN IN γ PER 24 HRS.
Franke and Fikentscher (25b)	10–30
van den Bergh et al. (96c)	10–100
Brugsch (9a, d)	4–50
Tropp and Siegler (93a)	18–110
Vigliani et al. (98d)	6–46
Vannotti (97a)	10–80
Carrié (11)	0–60
Lageder (52)	0–100
Thiel (92a)	0–100
Dobriner et al. (16n)	41–120
Watson et al. (100s)	14–99
Mason and Nesbitt (58a)	0–100
Franke and Litzner (25c)	3–7 (per 100 cc.)

ranging from 1 to 8 γ per 100 cc. (25c; 100s; w), and from 0 to 120 γ per day. The results of a number of investigators, as obtained with a variety of methods (vide infra), are shown in table 1.

The methods of determination have been, in the main, fluorimetric, depending upon the characteristic red fluorescence of the porphyrins in Wood's light. In accordance with Stoke's rule, the excited fluorescent light is of longer wave length than the exciting light which is absorbed (5a). For the coproporphyrins in 1 per cent HCl the exciting light is 4010Å (max.) (80f), while in 5, 10, 25 per cent HCl, the exciting wave lengths are 4018, 4030, and 4058 Å (max.) respectively (79d). Fluorescence is excited to a lessor extent by light of 5480 Å (max.). The fluorescence spectral band is found between 6200–6300 Å, maximum at 6250 Å

(51b.). The fluorescence intensity has been measured in various ways including the Pulfrich stufenphotometer (22; 25b; 97c; 98d), or fluorophotometers such as the Pfaltz-Bauer (58a) Klett (100x), or specially constructed types (63; 96a). Spectrometric (16f; 78f, 79c), and colorimetric methods (16n; 92a, b) have also been employed. Fluorescence methods are considerably more sensitive and require smaller starting volumes of urine.

In the study of isomer distribution in 41 normal individuals, as referred to in the above, the 24 hour quantity of total corproporphyrin was likewise determined in each instance. The range was from 14–99 γ. The statistical significance of these findings is discussed in more detail elsewhere (100s). On the basis of this study, it is believed that the upper limit of normal is 100 γ per 24 hours. The principal sources of error contributing to higher values obtained with colorimetric methods, as reported by others (20; 92b) are thought to be the measurement of impurities which are included due to insufficient preliminary purification. Thus it has been noted repeatedly that unless the porphyrin is taken back and forth twice between ether and HCl, unduly high values result, especially when the method depends upon absorption. Inadequate purification leads to unduly high values with colorimetric techniques and low values with those based on fluorimetry, in the former because of increased absorption of light, and in the latter because of masking of fluorescence, by impurities. Thus the method of Tropp and Siegler (93a) employs a continuous extraction apparatus which entails prolonged heating in boiling ether, together with any other impurities extracted from the urine. Experience[3] with continuous ether extraction has revealed consistent loss, as contrasted directly with repeated extraction in a separatory funnel. Nevertheless, Tropp and Siegler's range of normal values is slightly higher than ours, as noted in table 1. The reason for this is not apparent.

Poulikakos and Tropp (69) found little or no relation of the per diem amount of porphyrin to the volume of urine, in normal individuals. With increased excretion of porphyrin, however, relative increases were observed with larger volumes of urine.

Liver disease and jaundice. Garrod (27a, b, c) emphasized the occurrence of increased amounts of porphyrin, believed by him to be hematoporphyrin, in the urine in cases of liver disease. This has been confirmed by many subsequent investigators (9d; 16a; 24o; 34b; 65; 96c; 100b, t, u). Fischer and Zerweck (24o) showed that it was coproporphyrin which was increased, but did not determine the preponderant isomer type. van den Bergh (96c) found that coproporphyrin appeared in the serum and was increased in amount in the urine, both in cases of liver disease and of mechanical or obstructive jaundice. He postulated that this was due simply to a faulty excretion by the liver of the coproporphyrin normally present in the bile (96c). As will be discussed in more detail later, van den Bergh (96c) believed that this coproporphyrin is derived from the protoporphyrin of the erythrocytes, the conversion from one to the other being regarded as a normal function of the liver. Although later studies (77; 98a;

[3] Unpublished.

100n) have not supported this belief, the finding has been generally confirmed that coproporphyrin is regularly increased in the urine of cases of regurgitation jaundice whether due to diffuse liver damage or to biliary obstruction (9d; 16a, i; 65; 100b, t, u). Thus a recent study (100o, t, u) has revealed a significant elevation, usually from two to four times the upper limit of normal, in a series of cases of mechanical jaundice, and of jaundice due to hepatitis and cirrhosis. van den Bergh (96c) observed somewhat higher values in mechanical than in parenchymal jaundice, but did not attempt to employ the urinary coproporphyrin value in a differential diagnostic way. The data recently reported (100o) are

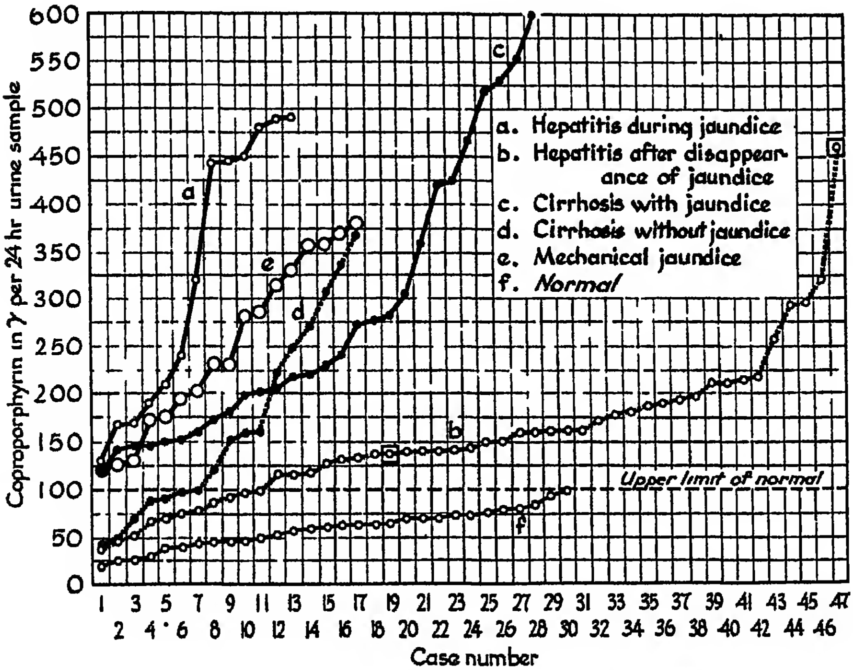


Fig 2. Total coproporphyrin in cases of hepatitis and cirrhosis. Case 19 in group b was studied about 18 months after the acute attack of infectious jaundice; case 47 in this group about 3 months afterward

shown in figure 2. It is evident from these and additional data which are discussed in separate communications (100t, u) that the total coproporphyrin determination has relatively little value in the differential diagnosis of jaundice.

Increase of the urinary coproporphyrin has been referred to frequently as a sensitive index of liver functional impairment (25a, b; 65; 93a). After alcohol ingestion, Franke (25a, b) observed prompt sharp increases which averaged 100 per cent, following one liter of beer or 90 cc. of cognac. Significant and rather uniform increases were also recorded by Franke in cases of "catarrhal" jaundice, subacute atrophy of the liver, cirrhosis of the liver, mechanical jaundice, and passive congestion of the liver. Franke noted generally higher values in

parenchymal than in mechanical jaundice. He did not regard the high values observed in 45 cases of lead poisoning as related to liver injury, as he found no evidence of the latter from the standpoint of other investigations. This matter will be considered in more detail later.

The first determination of isomer type in an instance of liver disease was carried out on urine from a case of cirrhosis of the liver which followed the use of cinchophen (100e). From the standpoint of the effect of chemical toxicity on porphyrin metabolism (vide infra), it may be noted, however, that a considerable interval of time had elapsed in this case, between the ingestion of the cinchophen and the time of isolation of the porphyrin. The latter proved to be the type I isomer. Subsequently, Dobriner (16a) isolated coproporphyrin I from the urines of a number of cases of various diseases, including atrophic cirrhosis, catarrhal jaundice, obstructive jaundice, hemolytic jaundice, Hodgkin's disease, and febrile states such as lung abscess and lobar pneumonia. Coproporphyrin I was likewise isolated from the urine in cases of fever due to lung abscess, and in hemolytic jaundice, by Watson (100b) who also found it in the bile (100f), and believed that the increase in these conditions was related to impaired hepatic excretory function. In Dobriner's study, an excess of coproporphyrin I was found in all of nine cases of catarrhal jaundice. These were presumably instances of infectious hepatitis. More recently, uniform increases of the type I isomer have been noted in the urine during the icteric stage in 57 cases of this disease (100o, t). In several of the cases of this series, the 24 hour amount exceeded 400 gamma. Of greater interest, however, was the finding of an increased urinary coproporphyrin for varying periods, often many months, after the disappearance of jaundice (see fig. 2). This persistence, together with collateral evidence has been interpreted as an indication of the presence of residual or chronic hepatitis. Similar slow decreases during convalescence of "catarrhal jaundice" had been observed previously (9d; 16a; 93a). In cases of cirrhosis following infectious hepatitis the excessive coproporphyrin has likewise been found to consist of the type I isomer (100o, t). This finding is in striking contrast to that in the majority of cases of cirrhosis in chronic alcoholics, in which the increase is due to the type III isomer (100u). Dobriner had previously reported an excess of type I in 6 cases of "atrophic cirrhosis," and of type III in one such case, but without stating which, if any of these instances, were chronic alcoholics.

The effect of alcohol on porphyrin metabolism has not yet received sufficient study. This problem will be dealt with now because of its close relation to that of hepatic cirrhosis, rather than in the later section dealing with chemical toxicity. Franke's finding of a sharp increase of coproporphyrinuria in acute experiments, has already been cited. This was confirmed by Brugsch and Keys (9f) who also noted increases in chronic alcoholism and isolated the type I isomer in two cases and the type III in one. Watson (100i) isolated coproporphyrin III from the urine in alcoholic pellagra (vide infra). In a study which is still in progress and which will be reported in detail elsewhere (90), it has been found that the great majority of chronic alcoholics excrete excesses of type III isomer while drinking

but usually exhibit a normal porphyrin excretion within a short time, often indeed a few days after complete abstinence from alcohol. Thus it is of considerable interest that an increased coproporphyrin III excretion persists in many instances of cirrhosis, in chronic alcoholics, long after the ingestion of alcohol has stopped. It appears, in other words, that some peculiar factor is operative in these cases which may relate not only to the persistence of the type III coproporphyrinuria, but also to the development of cirrhosis. Illustrative data are presented in separate communications (90; 100u).

Dobriner (16a) found coproporphyrin III in a case of hemochromatosis. In unexplained contrast to this finding, an excess of coproporphyrin I has recently been observed in the urine of two cases of hemochromatosis (100u). This discrepancy will be referred to again in the following.

Blood diseases. Fischer and Zerweck (24m) identified coproporphyrin in the urine of *pernicious anemia* patients. Garrod (27b) had previously referred to this as hematoporphyrin. Günther (34b), however, failed to observe increased "hematoporphyrin" in the urine in cases of pernicious anemia. The coproporphyrin in both urine and feces in this disease has been shown to be the type I isomer (16h; 98c; 100f). Inconstant increases in the urine of moderate degree have been discussed by Dobriner and Rhoads (16g). The same is true in *hemolytic anemia*, which, like pernicious anemia in relapse, is characterized by considerable increases of coproporphyrin I in the feces. There is some evidence that the latter is related to the rate of erythropoiesis (16d, h, m; 80f; 100f) (vide infra). The inconstant and variable increases in the urine in both of these diseases may well be analogous to the variable urobilinogenuria which is believed to relate to hepatocellular functional impairment (100a, f).

Vannotti (97c) was the first to report an increase in urinary porphyrin in a case of *aplastic anemia.* Dobriner, Rhoads and Hummel (16j), studied six cases of refractory or aplastic anemia in four of which an increase was found to be composed of a mixture of coproporphyrin I and III. In two additional cases, the presence of the type III isomer was suggested, but could not be proven. Dobriner and his co-workers considered the possibility that the increase of coproporphyrin III in these cases indicated a toxic or extrinsic chemical factor in their etiology. The relation of chemicals and heavy metals to type III coproporphyrinuria will be discussed in more detail in a later section. Recent studies of urinary coproporphyrin excretion in 12 cases of refractory or aplastic anemia, carried out by means of the differential precipitation technique (80b, e) have revealed a range of total coproporphyrin of from 119–363 γ per 24 hours, of which 56–92 per cent was the type III isomer. These data will be presented in detail in a separate communication (100v).

Inconstant mild increases of urinary coproporphyrin in leukemia have been reported in the literature and are discussed by Dobriner (16b, i). Dobriner (16b) found increases in cases of Hodgkin's disease ranging from 200 to 600 micrograms in febrile cases, shortly before death. This was ascertained to be the type I isomer in 6 instances. It was therefore of particular interest that in the course of a more recent study (100v), relatively large amounts of copro-

porphyrin III were encountered in three cases of Hodgkin's disease, while in two others, the excess was the type I isomer. This difference has not been explained; it will be referred to again in the discussion of origin and significance of the urinary coproporphyrins.

Hemorrhage and the effect of blood in the gastrointestinal tract. Duesberg (17a), studying hemorrhage in rabbits, did not observe any increase in urinary porphyrins. Dobriner and Rhoads (16h) found an increase in total coproporphyrin excretion in dogs, after bleeding from the jugular vein; the percentage of this increase, if any, excreted in the urine, was not stated. Dobriner's interpretation of this finding will be discussed later.

Increases of urinary porphyrin have repeatedly been ascribed to the presence of blood in the intestinal tract (6; 9a, b; 24h; 79c; 86a; 97c). The inconstant character of these increases induced Kaunitz (46) to administer hemoglobin, in 20 gram quantities, to normal individuals and to patients with various disorders, including liver disease. It was found that negligible, if any, increase occurred in the former, but that significant elevations were common in the latter group, especially in the cases with liver disease. Kaunitz interpreted this to indicate a formation of the type III isomer in the intestinal tract followed by partial absorption into the portal circulation with disposition by the normal liver, so that little or none was allowed to enter the general circulation and appear in the urine, while in the case of the damaged liver variable fractions gained access to the urine. The study of Jakob (42) offered indirect support for this belief. According to this study certain bacteria are capable of elaborating coproporphyrin III from hemoglobin, in vitro. Kämmerer (43a) had previously demonstrated the bacterial formation of protoporphyrin from hemoglobin, in vitro, but he believed that this transition required a symbiotic growth of anaerobes with E. coli, while others have shown that symbiosis is unnecessary, and that in the presence of hemoglobin, pure cultures of E. coli are capable of the elaboration of protoporphyrin (38a) and of coproporphyrin III (42).

The question of the relation of blood in the gastrointestinal tract, and of the bacterial flora of the intestine, to the urinary coproporphyrin, under normal and pathological conditions, has been studied in some detail in this laboratory (53a, b). Jakob's findings have been confirmed including the observation that certain strains of bacteria can synthesize coproporphyrin III in the absence of any demonstrable heme pigments or porphyrins, but in the presence of protein. It was not elaborated in synthetic media of the Gladstone type. In this connection it may be noted that Coulter and Stone (13) observed formation of coproporphyrin by diphtheria bacilli in bouillon cultures. The rate of formation appeared to be proportional to the toxin production. The isomer type was not determined.

Nevertheless, the recent study just referred to failed to support Kaunitz' view that hemoglobin in the intestinal tract causes increased coproporphyrinuria, even in the presence of liver functional impairment. No significant increases were observed in several instances of outspoken liver disease, following feeding of hemoglobin in amounts as large or larger than those which Kaunitz gave. It

is believed of even greater significance that there was repeated failure of absorption of coproporphyrin III given by mouth, after administration of 3–5 mgm., together with lack of increase of the coproporphyrinuria in such experiments. Thus it would appear that while the bacterial flora of the colon may produce small amounts of coproporphyrin III even in the absence of hemoglobin, the presence of this porphyrin in the bowel bears little or no relation to the amount in the urine. The results of this study, in fact, strongly indicate that the coproporphyrin III of human urine is a truly endogenous metabolite. Whether this applies as well to the type I isomer is not certain. This problem will be considered later (vide infra), but it may be noted at this point, that in one experiment in which 3 mg. of coproporphyrin I was administered orally, no increase was observed in the urine and approximately the amount given was recovered in the feces.

There is no evidence that chlorophyll in the diet contributes to the coproporphyrin of the urine. Brugsch (9d, g) has discussed this problem in detail.

Fever and infectious disease. The finding of coproporphyrin I in febrile states associated with pneumonia (16a) and lung abscess (16a; 100e) has already been referred to, and the possibility was considered that these increases were related to hepatic functional impairment.

More recently Cartwright and co-workers (11b) have observed increased coproporphyrinuria regularly in association with the anemia of pyogenic infection. The isomer type was not stated.

The uniform increase of coproporphyrin I in the urine of cases of infectious hepatitis during the icteric stage of the disease, and in many instances persisting after disappearance of jaundice, has also been mentioned. It was therefore an unexpected finding of much interest that the urine of cases of *acute poliomyelitis* regularly contained significant increases of type III isomer (100r). In general, larger amounts of total coproporphyrin, ranging up to 1000 gamma per day with 60–90 per cent of coproporphyrin III, were encountered in the more seriously ill patients, especially the bulbar type, but a strict correlation with severity or type has not been determined. Of 64 cases of the disease, 56 exhibited amounts of total coproporphyrin in excess of 100 gamma. In some instances in which the total coproporphyrin value was normal, the percentage of the type III isomer was abnormally high. Figure 3 shows the total coproporphyrin values and isomer distribution in 51 cases, as compared with the normal range. In a number of these instances crystalline coproporphyrin III methyl ester was isolated, thus confirming the results obtained by means of the differential precipitation technique (80b, e). The cause of this increase is unknown. Klüver (51a, b) established the presence of small amounts of coproporphyrin in the white matter of the central nervous system of warm blooded animals, and identified it tentatively as the type III isomer (51c). Based on studies of bacterial metabolism, Granick and Gilder (31) suggest that coproporphyrin may function as a regulator for the rate of oxygen consumption by living cells. It remains to be determined whether the coproporphyrin of the central nervous system has such a rôle. Studies by Chu and Watson (12)

have also tentatively identified the porphyrin as coproporphyrin III and have established the concentration at about 0.02–0.03 gamma per gram of human brain. It was found impossible to increase the concentration in rabbit brains by poisoning the animals with lead, in spite of the fact that a 10 or 20 fold increase of urinary coproporphyrin III was effected (12). Nor was it possible to achieve any increase in concentration in dogs' brains by injecting relatively large amounts of coproporphyrin III into the internal carotid artery one-half hour before the animals were killed and the brains examined (12). These observations render much more difficult the correlation of the above mentioned urinary findings in poliomyelitis with the central nervous system coproporphyrin; nevertheless, there may be a relationship and this possibility is receiving further study at the present time.

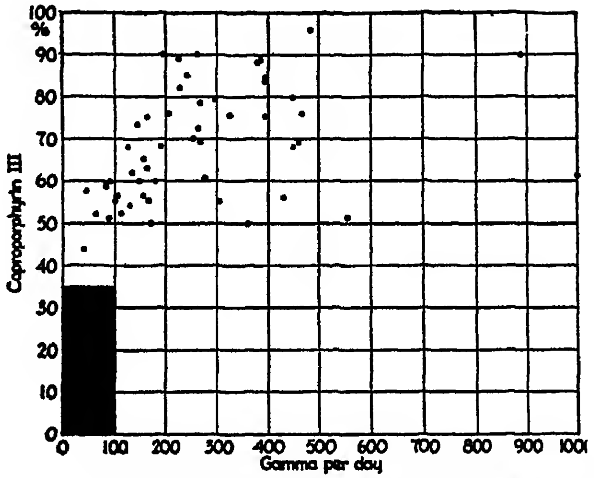


Fig. 3. Total coproporphyrin and per cent of type III isomer in 64 cases of poliomyelitis. The shaded area in the lower left corner represents the normal range. From the paper "Coproporphyrnuria (type III) in acute poliomyelitis," by C. J. Watson, Wm. Schulze, Violet Hawkinson, and A. B. Baker, Proc. Soc. Exper. Biol. and Med., **64**: 73, 1947. Reproduced by permission of the publishers.

The fact that two virus diseases such as infectious hepatitis and acute poliomyelitis are sharply differentiated on the basis of the urinary coproporphyrin isomer distribution, indicates that other virus diseases should be investigated in this regard. A study is in progress which will embrace both neurotropic virus affections such as encephalitis and Guillain-Barre' radiculitis, and those which are not neurotropic, such as atypical pneumonia, psittacosis and others. At the present time, the literature does not contain any data of this type. It is of interest, however, that in another disease of the brain, namely schizophrenia, Libowitzky and Scheid (55) report that the febrile episodes which are often observed, were characterized by an increased urinary coproporphyrin type I.

Kapp and Coburn (44a) described an increase of urinary coproporphyrin in acute rheumatic fever, thus giving emphasis and definition to the observations of

MacMunn on urohematin (56), and of Garrod (27a) and Stokvis (86b) on hematoporphyrin, many years earlier. Kapp and Coburn were able to isolate coproporphyrin III from the pooled samples of a number of cases. Kapp (44b) noted that the coproporphyrin was excreted in considerable part as the zinc complex. In several instances "sudden fatalities were preceded by intense porphyrinuria." In addition Kapp and Coburn state "the excretion of abnormally large amounts of porphyrin in acute rheumatism bore no relation to fever, haematuria, or the use of therapeutic measures per se" No mention was made in these papers of the use of salicylates or amidopyrine, although Dobriner and Rhoads (16i) state that Kapp and Coburn's patients were receiving these drugs at the time of the porphyrin isolation. Brownlee (8a) showed that amidopyrine, amongst other drugs of the coal tar group, produced coproporphyrinuria type III in rats. So far as can be determined, the effect of salicylates on porphyrin excretion has not been investigated. Watson and Hawkinson[4] have found moderate and variable increases of urinary coproporphyrin in acute rheumatic fever. These increases, however, were composed in all instances of the type I, rather than the type III isomer. The possibility was considered that the basis of the increase was a febrile or toxic effect on the liver. The administration of salicylates was without definite effect on the porphyrin excretion, no increases of type III isomer being observed.

Porphyria and idiopathic coproporphyrinuria. As noted at the outset, the entity of coproporphyrin was first established by H. Fischer who isolated it, together with uroporphyrin, from the urine of Petry, the celebrated case of porphyria (24a). This coproporphyrin was subsequently classified as the type I isomer (24c). Coproporphyrin III was likewise first obtained from the urine of a case of porphyria of the congenital, or light sensitizing variety (24e, f; 96a). Experience has revealed, however, that this latter finding is somewhat exceptional and that it is more common to find the type III isomer in the urine of the so-called acute type, i.e., that which is characterized by abdominal or nervous manifestations, while the type I isomer is most often encountered in the light sensitive variety, such as represented by the case Petry. The idea has even been advanced (62; 78a) that a strict chemical separation of the two types might be made on a chemical basis, but it is evident that there are too many exceptions, instances in which coproporphyrin I was isolated from the abdominal or nervous type (11a; 24j; 101), and coproporphyrin III from the light sensitive type, both in human cases (24f; 96d), and in cattle (73a, e), or, as in certain cases, both isomers in approximately equal amount (16b; 96a). Thus the case of van den Bergh and Grotepass (96a) was an instance in which photosensitivity was prominent, and which is also of much interest because of coproporphyrinemia and the presence of large amounts of the two coproporphyrin isomers in the feces, without a definite increase in the urine. In this case, Grotepass and Defalque (33c) were able to achieve isolation only by collection and concentration on infusorial earth, over a long period of time. Taylor and

[4] Unpublished.

co-workers (91) have recently isolated both isomers I and III from a case of "chronic" porphyria in which the prominent clinical features were abdominal pain and skin lesions due to photosensitivity. As mentioned at the outset, the term porphyria connotes the formation of uroporphyrin, for which there was no evidence in either of these last two cases. At present the question cannot be answered as to whether such cases differ fundamentally from porphyria or represent simply an unusual variation of the disease. As noted in the following it is even more difficult to classify certain instances in which large excesses of coproporphyrin are excreted without co-existence of any of the usual symptoms of porphyria.

The separation of the two main clinical types of porphyria on the basis of isomer I or III excretion became still further complicated with the finding that the Waldenström porphyrin, which had come to be regarded as uroporphyrin III (11a; 60b; 99a, b, c) is actually a complex mixture of uroporphyrin I and of a type III porphyrin which may have seven rather than eight carboxyl groups (32c; 100q). The Waldenström type porphyrin has only been encountered in the urines of cases of the intermittent, acute porphyria and may undoubtedly be regarded as a distinctive feature of this form of the disease, although not as an isomer of uroporphyrin I.

Waldenström (99c) stated that in the acute type of porphyria little or no coproporphyrin was to be found in the urine and that for this reason he had not isolated it. In two cases in which uroporphyrin was isolated from the urine, Waldenström obtained coproporphyrin I from the feces. H. Fischer and Libowitzky (24j) had previously found both uro- and coproporphyrin I in the urine of a case of acute porphyria. In the cases studied by Mertens (60b) and Bingel (4a), the daily amount of uroporphyrin in the urine ranged from 50–100 mgm., while the coproporphyrin approximated 1 mgm. Recently Brunsting and Mason (10b) described a case of porphyria of mixed type characterized by epidermolysis bullosa, pigmentation of the skin and neuropathy. During one 24 hour period this patient excreted 11.4 mgm. of uroporphyrin and 2.0 mgm. of coproporphyrin. Uroporphyrin I was identified, but the isomer type of the coproporphyrin was not determined. While the concentration of coproporphyrin in the urine of cases of intermittent acute porphyria is thus comparatively slight, it usually suffices to permit isolation from relatively small volumes. Representative data from a number of cases are given in table 2. Actually a number of other cases studied exhibited equal or greater concentrations of coproporphyrin but in these instances it was not quantitated, the main attention being given to isolation of the Waldenström porphyrin (32c; 100q).

In certain cases, as for example no. 9 in table 2, the amount of coproporphyrin considerably exceeded that of uroporphyrin. The reason for this is not clear. The data in this case, over a considerable period of time, are shown in figure 4. As may be noted, this patient suffered a spontaneous relapse while in the hospital, for which no precipitating factor was ascertained. During this relapse it is seen that the urine was characterized by an increase of all of the abnormal

elements, including the porphobilinogen and porphobilin, uro- and copro-porphyrins. It may be noted, too, that the amount of zinc complex was relatively larger during the relapse. The significance of the combination of uro- and coproporphyrin with zinc, is unknown. Günther (34a) recognized that the urinary porphyrins in cases of porphyria were often partly combined with a metal. Derrien and co-workers (15a, b) described zinc uroporphyrinuria and even ascribed a protective or detoxifying action of the zinc. So far as can be determined this suggestion has not received further attention. Watson and his associates (100k, o, q) have likewise observed the zinc complex of Walden-

TABLE 2

Relative amounts of copro- and uroporphyrin in cases of intermittent acute porphyria

PATIENT NUMBER	COPROPORPHYRIN IN γ/24 HRS	PREDOMINANT TYPE ISOMER	UROPORPHYRIN IN γ/24 HOURS
1	144	(III)	708
2	378	(III)	5,250
3	156	(I)	2,340
4	460	(I)	83,000
	554		147,000
	801		87,000
5	160		508
6	182		
	250		
	312		
	300		
7	307	(I, III)	
	406		34,000
8*	256		120
	240	(III)	
9 (see fig. 4)	401	(III)	160
	484		110
	2,582		412
	1,306		307
	150		75
10*	182	(III)	61

* Latent cases; mother and aunt, respectively, of case 9.

ström's uroporphyrin as a regular component of the urine in cases of intermittent acute porphyria. The data shown in figure 4 reveal that a considerable fraction of the coproporphyrin was likewise complexed with zinc. The possible contribution of soft glass to this complex was excluded by making the collection in Pyrex bottles only. The occurrence of the zinc complex of coproporphyrin is not peculiar to porphyria; Mertens (60a) observed it in the urine in cases of lead poisoning, a finding which we can confirm. Thus in one instance the 24 hour urine contained 1686 γ as the metal complex and 1107 γ as free coproporphyrin. The absorption spectrum and HCl numbers have repeatedly been noted to correspond with those of the zinc complex (24k; 100q).

Dobriner and co-workers (161) isolated a natural ester of coproporphyrin I from a case of congenital porphyria. This melted at 251°C. The HCl number was 0.3–0.5 per cent, from which concentration the compound was readily extracted by $CHCl_3$. While the esterifying group could not be identified, it was thought to be lipid in character. So far as can be determined this has not been encountered in other cases.

As noted in figure 4, porphobilinogen was present in the urine in association with uro- and coproporphyrin. The term porphobilinogen refers to a chromogen which exhibits an Ehrlich aldehyde reaction, and which is believed to be a precursor of uroporphyrin and porphobilin (99c, g). The latter is a brown pigment not possessing porphyrin characteristics and formerly referred to as "urofuscin"

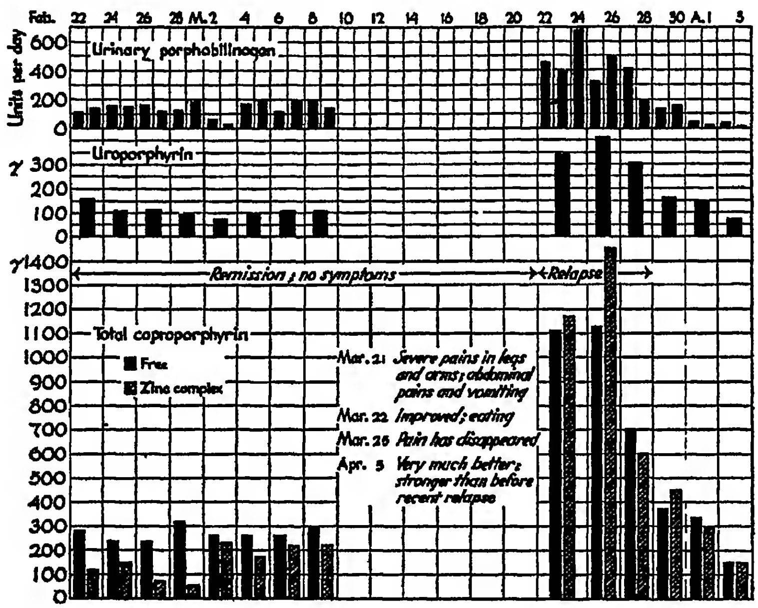


Fig. 4. Acute porphyria; excretion of porphobilinogen and of uro- and coproporphyrin, both free and zinc complex, during remission and relapse. M. ♀ C. 25 (case 9 in table 2).

(99c). It has long been known that the urines of cases of "acute" porphyria often exhibit an Ehrlich reaction, although whether due to urobilinogen or porphobilinogen was not determined in the past, since a method for differentiating these chromogens has only lately been described (99c; 100y). Insofar as the present subject is concerned, the question arises as to whether coproporphyrin is likewise excreted at least in part as a porphobilinogen. It is quite clear that even marked increases of urinary coproporphyrin are not associated with a porphobilinogen-Ehrlich reaction (100y), except in cases of porphyria in which uroporphyrin is also present. We have observed but two instances in which questionable reactions were encountered, and we have not found any proven examples in the literature. This does not, of course, exclude the possibility

that some of the coproporphyrin is excreted as a chromogen, but if so it does not have the behavior of porphobilinogen as found in the urine of porphyria cases, such as shown in figure 4.

The presence of large excesses of coproporphyrin, or of uroporphyrin, in the urine, is not correlated with the appearance of the corresponding porphyrins in the blood. In some instances uro- and/or coproporphyrin may be noted. Schumm (79a) observed uroporphyrin in a case of congenital porphyria in an approximate concentration of 1 mgm. per 100 cc. The urine in this case contained both uro- and coproporphyrin I. H. Fischer and Zerweck (24m) later reported the presence of coproporphyrin in the same case; in a separate report in the same year, however, they described uroporphyrinemia in this case (24o). In other cases of porphyria, both congenital (light sensitive) and intermittent acute in type, coproporphyrin has been noted in the serum (40; 78d; 96a; 96d). We have noted it repeatedly in 3 cases of the latter type, and have not been able to detect it in 6 others.

In one case the amount of coproporphyrin in the serum was determined to be 20 γ per 100 cc., while at the same time, the 24 hour urine sample contained 300 γ of coproporphyrin, most of which was type III. This sample also contained 1800 γ of uroporphyrin of the Waldenström type melting at 258°C. and separating on the $CaCO_3$ column into uroporphyrin I methyl ester (284°C.), and the 208° porphyrin (32c.).

The protoporphyrin of the erythrocytes has not been increased in several of our cases of intermittent acute porphyria, even during relapses of the disease when large amounts of uro- and coproporphyrin were found in the urine. Roth (75a) on the other hand studied a case in which the erythrocytes contained 74 γ of protoporphyrin per 100 cc., the urine 79 γ of coproporphyrin in a 24 hour sample. This was classified as an instance of "chronic" porphyria (Günther), since abdominal pain was associated with light sensitivity. The presence of diarrhea and hypochromic anemia suggest, however, that the moderate increase of the erythrocyte protoporphyrin was due, not to porphyria, but to iron deficiency (100p); the value for the urinary coproporphyrin was within normal limits.

The problem of excessive coproporphyrinuria without known cause is of much interest. For the time being at least, it is believed preferable to classify such cases, in which no uroporphyrin is found, as *idiopathic coproporphyrinuria*, rather than as porphyria, reserving the latter term to indicate a constitutional fault. This is in agreement with the separation emphasized by Waldenström (99c). At the same time it is recognized that certain cases excreting increased coproporphyrin, without uroporphyrin, may also represent "inborn errors of metabolism" (27d). A particularly interesting example of marked coproporphyrinuria without known cause has been studied recently (100w). This individual had no symptoms whatever, the excessive urinary coproporphyrin being noted in the course of a routine study of a large number of workers engaged in a special project. The urinary coproporphyrin was consistently in the range of 1–6 mgm. per day, practically all of which was the type III isomer. The

feces contained larger amounts of the same porphyrin, ranging up to 70 mgm. per 100 grams. None was detected in the blood at any time. It was readily established that the bile was rich in coproporphyrin III, but whether due to an enterohepatic circulation of porphyrin, could not be established. It appeared unlikely that an abnormal bacterial flora in the colon constituted an etiologic factor, since the bowel movements were normal and there were no symptoms referable to the intestinal tract. It was of interest that this man was a chemist who had worked in the past with heavy metals, especially lead and mercury. While there were no evidences of metal poisoning, the possibility was considered that a marked over production of coproporphyrin III may have been related to the effects of heavy metals or other chemicals (vide infra) in an individual of peculiar constitution.

Several cases have been studied (100x) in which intermittent abdominal pain was associated with moderate increases of urinary coproporphyrin, ranging from 120–600 γ per day, with an increased percentage of the type III isomer. Nothing was found in these cases to explain the increases. The fecal flora studied in certain of the cases did not appear abnormal. Nor were the symptoms or porphyrin excretion affected by a period on the Kāmmerer diet (43b). This meat free diet is low in heme and porphyrin and was used by Kāmmerer and Meyer in a similar case which they studied. Although they believed that the addition of liver or meat to this diet caused a significant increase of coproporphyrinuria, the data given are not convincing.

Barber and Howitt (1) reported several instances of photosensitivity in children, in which an increased fecal porphyrin was observed. It was not at all clear, however, that the porphyrin was in any way related to the photosensitivity. The type of porphyrin was not determined either in feces or urine and the statement that porphyrin was increased in the feces was based solely upon spectroscopic studies. Urbach (95a, b) also associated spasticity of the colon and dermatitis, especially solar eczema, with increased porphyrin formation in the bowel which was ascribed to an abnormal bacterial flora. Quantitative methods were not employed in this study. Alteration of the flora toward the normal, ascribed to elimination of animal protein from the diet, was said to result in marked improvement. Haranghy (36) reported an interesting case characterized by abdominal colic and circulatory disturbances. The urinary porphyrin was not increased but there was a marked increase in fecal porphyrin (type not mentioned). This was believed dependent on an abnormal flora which was almost wholly gram-negative It may be noted, however, that there were bleeding intestinal ulcers in the case, which may well have been significant from the standpoint of porphyrin formation.

The study of Reitlinger and Klee (72) is of interest with respect to the possible relation of coproporphyrin to abdominal colic, although it does not aid in determining the origin of the porphyrin. This study showed that very dilute solutions of coproporphyrin I instilled into exposed loops of rabbit intestine caused prolonged spastic contraction which was refractory to atropine.

The effect of chemicals and metals. In this group the effect of lead poisoning

was the first to receive study, increases in urinary porphyrin being observed by Garrod (27a, c) and Stokvis (86a) in human cases, and by Stokvis in experimental lead poisoning in rabbits (86a). The latter investigator believed that the porphyrin was derived from hemoglobin in the intestinal tract, which in turn was liberated by virtue of mucosal hemorrhages. In later studies (vide infra) it has been shown repeatedly that coproporphyrin III is the porphyrin which appears in the urine in excess, and, as noted previously, the evidence for a derivation of coproporphyrinuria III from blood in the intestinal tract is very dubious; furthermore, intestinal bleeding is not a feature of human lead poisoning. Stokvis' view has not been supported by any subsequent studies although it was subscribed to by Brugsch (9d). In several instances the feces have contained coproporphyrin I at a time when the urine yielded marked excesses of the type III isomer (98c; 100e).

Schumm (79b) was the first to identify coproporphyrin in the urine of cases of lead poisoning, and Grotepass (33b) first showed that it is the type III isomer, a finding which Fischer and Duesberg (24f) reported in the same year with respect to experimental lead poisoning in rabbits. Subsequent studies (16h; 60a; 98e; 100e) have extended and confirmed these observations with the result that coproporphyrinuria III has come to be recognized as a regular feature of lead intoxication. Günther's belief (34a) that this was of diagnostic value has been amply confirmed by later work (25c; 45b; 67; 78e) which has emphasized its value in industry. Franke and Litzner (25c) found a range of 12 to 181 γ per 100 cc. of urine in 43 workers, as compared with a normal of 3–7; 17 of the 43 had concentrations greater than 50 γ per 100 cc., and most of these exhibited other evidence of lead poisoning, including basophilic stippling. The experience of various investigators is shown in table 3.

The amounts noted in table 3 are insufficient, as a rule, to cause a red urine. This is, in fact, but rarely noted in lead poisoning. A case in which the urine was "port wine" color was observed by Crawford (14).

Knowledge of the effect of other metals on the urinary coproporphyrin is much more fragmentary. Considerable increases following mercury have been recorded by Vannotti (97b) and by Vigliani and Libowitzky (98c), but in both of these cases there were confusing complications, viz., a hematuria and hemoglobinuria in the former and a cirrhosis of the liver with chronic alcoholism in the latter. In Vannotti's case, there was light sensitivity and both uro- and coproporphyrin I were isolated, the daily output of the latter during the period of manifest illness ranging from 4 to 12 mgm. Two years later there was still light sensitivity and the coproporphyrin excretion ranged between 200–1300 γ per day. The uroporphyrin was found for a short time only during the acute poisoning and brings up the question as to whether this was a case of latent porphyria in which mercury was a precipitating factor.

Insofar as the case of Vigliani and Libowitzky is concerned, it has already been noted that coproporphyrin III is usually present in excess in cases of cirrhosis in chronic alcoholics, and thus it would be difficult to ascribe the excess in this case to mercury.

Hoerburger and Fink (39b) isolated coproporphyrin III from the urine of a case of salvarsan toxicity which had been studied by Schreus (78c). Subsequent studies (10c; 81) have revealed frequent significant increases in patients receiving salvarsan, more particularly in cases having exfoliative dermatitis (10c).

We have recently investigated a case of chronic arsenic poisoning characterized by typical hyperkeratoses, peripheral neuritis, and anemia. Arsenic was found in marked excess in the hair and urine, and, although the source was never proved, there was reason to believe that the patient had been poisoned intentionally. The urinary coproporphyrin excretion ranged from 1 to 3.2 mgm. per day. Since this patient had worked as a filling station attendant, at times siphoning tetraethyl lead gasoline and handling batteries, the possibility was considered that he also had lead poisoning. The presence of severe anemia and basophilic stippling were somewhat more suggestive of lead than of arsenic poisoning. The erythrocyte protoporphyrin in this case ranged from 200–400 γ

TABLE 3

Urinary coproporphyrin excretion in humans with lead poisoning

AUTHOR	PER DIEM EXCRETION*	PREDOMINANT ISOMER TYPE
Grotepass (33a)	To 8.0 milligrams	Copro-III
Schreus and Carrié (78e)	To 1.8 milligrams	Copro-III
Roth (75b)	To 1.4 milligrams	Not determined
Watson (100e)	To 1.0 milligrams	Copro-III
Vigliani and Angeleri (98b)	To 3.4 milligrams	Not determined
Mertens (60a)	To 5.0 milligrams	Not determined
Vigiliani and Libowitzky (98c)	To 2.0 milligrams	Copro-III
Vigliani and Waldenström (98e)	To 3.1 milligrams	Copro-III
Kark and Meiklejohn (45b)	To 1.2 milligrams	Not determined

* The values are not entirely comparable, as they are based in some instances on the yield of crystalline ester, and in others on fluorimetry.

per 100 cc. of erythrocytes, as compared with a normal range of 10–30 γ (100p). Data are not available with respect to the erythrocyte protoporphyrin in arsenic posisoning. Serum and urine lead determinations in this case were just at the upper limits of the normal ranges and were therefore not diagnostic.

Certain other investigations of the effects of metals on porphyrin excretion in animals may be mentioned. Thus Sumegi and Putnoky (89) observed that Pb, Hg, Bi, Cu, Zn, Fe, Au and Ag caused loss of weight in white rats, together with hyperfunction of the thyroid and porphyrinuria. Also mentioned were disturbances of liver and renal function. The intensity of these disturbances decreased in general in the order of the metals given. Thyroid hyperfunction and disturbance of porphyrin metabolism were greatest with Pb, Hg and Bi. These investigators found also that porphyrin alone, injected into rats, caused intense thyroid hyperfunction. Because of this they concluded that the porphyrin stimulated centers in the mid brain thus inducing a thyrotoxicosis. Obviously this study is of much interest and the results need to be confirmed.

Recently the effect on the urinary coproporphyrin of lead, uranium, thorium, beryllium, lanthanum, zirconium, mercury, and arsenic has been studied in both acute and chronic experiments on rabbits (80c). Insofar as lead is concerned the results of this study have confirmed those of previous investigators. Marked increases of type III isomer excretion, often persisting for a number of months, were noted to follow single injections of from 8–50 mgm. of lead (given as lead acetate) per kilo of body weight. Thorium, zirconium and lanthanum did not cause an increase except in animals which had received lead several months previously. Since this was associated with an increased excretion of lead, it was postulated that the mobilized lead rather than the newly administered metal of other type, was responsible for the increased coproporphyrin III. Uranium was found to cause a marked decrease in coproporphyrin excretion during the acute stage of poisoning. This was ascribed to renal injury. Smaller amounts were followed by a sustained, moderate increase. Injection of arsenic resulted in considerable increases of short duration.

A considerable literature has accumulated with respect to the relation of *sedatives* to porphyrin excretion and porphyria. The use of sulfonal and trional has often been followed by a disease which is indistinguishable from the intermittent acute porphyria of idiopathic type. Uroporphyrin has been isolated repeatedly from the urine. Ellinger and Riesser (19) noted a melting point of 255–257°C. which suggests a Waldenström type porphyrin such as usually encountered in acute porphyria; Dobriner, however, isolated uroporphyrin I from his case (16a). The findings in animal experiments have been somewhat confusing. Stokvis (86a) first noted an increased urinary porphyrin excretion in rabbits and dogs, following sulfonal. Fischer and Duesberg (24f) identified uroporphyrin spectroscopically, in urine from sulfonal poisoned rabbits. Waldenström and Wendt, in a later study (99d), found only coproporphyrin III. It was of interest in this study that dimethylsulfon-dimethylmethane, which differs from sulfonal only in having non-hypnotic methyl rather than hypnotic ethyl groups, was just as effective in increasing the urinary coproporphyrin. Neubauer (66) had been aware of the similar effect of both compounds many years earlier, although designating the porphyrin as hematoporphyrin. Waldenström (99c) has shown that trional can precipitate attacks of typical acute porphyria in members of the family of individuals in whom the disease had previously been diagnosed. This emphasizes the question, as yet not answered, whether sulfonal and trional porphyria represent merely a precipitation of a latent, constitutional disease.

Brugsch (9d, e) noted distinct, although not marked, increases of urinary coproporphyrin in patients receiving *paraldehyde*, *chloralhydrate*, *amylene hydrate*, and *morphine*, but not after therapeutic doses of *barbiturates*. According to Laubender and Monden (54) the urinary porphyrin excretion in rabbits is not increased by feeding of barbiturates. Duesberg (17b) and Haxthausen (38) both reported individual instances of acute porphyria following use of barbiturates. In the latter case phenobarbital had been taken, and it appeared that there was complete recovery following cessation of the drug. In Duesberg's

case uroporphyrin was tentatively identified. The possibility must be considered, as in the case of sulfonal and trional, that the barbiturate was only a precipitating factor in a latent, constitutional disorder. Eichler (18) and Brugsch (9d, e) believe that latent disturbance in porphyrin metabolism characterized only by mild nervous and mental disorders, may be exaggerated to the point of outspoken or fatal porphyria by the use of sedatives.

The effect of sulfonamides and other aromatic amines on the urinary coproporphyrin has been studied in some detail. Significant increases have often been observed following sulfanilamide administration to humans (73d; 84), especially where toxic reactions were encountered (10a, c; 21; 100m). Rimington and Hemmings (73c) noted marked increases in the urine and feces of rats receiving large amounts of sulfanilamide. This was true, also, with many other aromatic amines which they tested. Similar results were obtained by Brownlee (8a), who studied sulfanilamide, acetanilide, phenacetin, phenazone, amidopyrine, aspirin, and p-aminophenol. In both of these studies the excessive coproporphyrin was shown to be the type III isomer. It was postulated that this was derived from methemoglobin, the latter being formed by the oxidative action of the chemical (see Brownlee's paper (8a) for further discussion). This possibility will be considered later.

Penew and Tropp (68) reported increased urinary coproporphyrin excretion in humans following ether or nitrous oxide anesthesia, and in some instances after novocain (local) anesthesia. The increase in all of these cases was thought to be related to transitory liver injury, classified by them as a "serous inflammation." Tropp and Penew (93b) noted a large amount of coproporphyrin in the urine of an individual poisoned with illuminating gas. The isomer type was not mentioned in either of these papers.

The effect of carbon tetrachloride upon the urinary coproporphyrin has not been discussed in the literature. In a case of acute poisoning observed recently, a marked increase of the type III isomer in the urine was determined. This patient, however, was a moderate alcoholic so that the finding is somewhat more difficult to interpret. Thus it was found that the patient excreted 373 γ in 24 hours shortly after admission; 65 per cent of this was the type III isomer, as determined by the differential precipitation technique. Gillam and Kench (29) studied a case of severe methyl chloride poisoning characterized clinically by delirium, convulsions, and tremor. Marked increases of coproporphyrin III were isolated both from the urine and feces. A rapid decline in the amount occurred as the patient recovered. There was no anemia at any time.

The effect of *alcohol* on the urinary coproporphyrin has already been discussed in the section relating to liver disease, particularly the cirrhosis of chronic alcoholics.

Skin diseases and photosensitivity. The fundamental aspects of photosensitisation by the porphyrins have been discussed in detail elsewhere (5a, b; 37b, c; 87a, b; 97c), and will not be considered here. In the main, outspoken photosensitivity has been a feature of congenital porphyria (uroporphyrin excretion) and has been noted but rarely in cases where coproporphyrin alone was observed

(vide infra). This is in accord with the observation of Fischer (24c) that in warm blooded animals, the degree of photosensitizing ability of the naturally occurring porphyrins is correlated with the number of carboxyl groups. Thus, uroporphyrin with eight is the most effective, coproporphyrin with four less so, and proto- and deuteroporphyrins with two each, not at all. On the other hand, the artificial hematoporphyrin, which also has but two carboxyl groups, is rather strongly light sensitizing both for man (24o; 61; 88) and animals (4b; 26; 28; 37b, c, b). Dobriner and Rhoads (16i) give the order of photosensitizing effect, as judged by the severity of "light shock" (24c) produced in white mice, as follows: "hematoporphyrin 9 (type III), uroporphyrin I, deuteroporphyrin 9 (type III), coproporphyrin I, and protoporphyrin 9 (type III). Coproporphyrin type III seems to be but slightly toxic and uroporphyrin 'type III' non toxic." Bingel (4b) also reported that coproporphyrin III is less sensitizing than I. Gildemeister (28) found that the uroporphyrins combine with the serum albumin fraction, but that coproporphyrins I and III do not. This is of obvious interest when it is remembered that blood serum has been reported to be protective against photosensitization (71; 82). Shibuya (82) found, however, that serum was also protective against the photosensitizing effect of coproporphyrin. Fischer and Zerweck (24m) had previously reported a protective action on the part of the brown pigment (urochrome) of normal and pathological urines, a compound which they believe was derived from protein. Mertens (60b) correlated the chemical findings in the urine with the presence or absence of photosensitivity, in the cases of porphyria reported up to the time of her paper (1937). All of these but two, i.e., the cases of van den Bergh (96d), and of Fischer and co-workers (24l), excreted uroporphyrin. The two exceptions, interestingly enough, both excreted coproporphyrin III, a finding in some conflict with the experimental results already referred to. As mentioned previously, the case of chronic porphyria recently reported by Taylor and co-workers (91) exhibited marked photosensitivity, yet the urine was said not to contain uroporphyrin. Both coproporphyrin isomers were identified. The case of Grotepass and Defalque (33c) appears to have been of the same type. The relation of coproporphyrin to the photosensitivity in these and other disease states is not at all clear. The case of idiopathic coproporphyrinuria cited earlier, in which unusually large amounts of the type III isomer were excreted, yet without manifestations of any kind, attests to the difficulty of the problem. Photosensitivity in certain skin diseases has been linked by certain investigators with a secondary coproporphyrinuria. Thus it has been suggested repeatedly that the dermatitis of pellagra is related to photosensitization by coproporphyrin (2; 3; 20; 59; 85). The methods used in these studies were not specific for porphyrin, and in fact it is clear that the color measured by the technique of Ellinger and Dojmi (20) is mainly due to formation of urorosein from indolacetic acid, rather than to porphyrin (100i, j, l). In alcoholic pellagra considerably increased amounts of coproporphyrin have been encountered in the urine (16k; 74; 100i). This has been identified as coproporphyrin III and has been noted to decline after nicotinic acid therapy and abstinence from alcohol (16k; 100i). As ob-

served earlier, however, this decline is quite usual in the ordinary alcoholic who stops drinking and resumes eating. Significant increases of coproporphyrin have not been found in endemic pellagra, either in urine (45a; 100i, l) or feces 100l). Ralsky (70) reported an increased porphyrinuria in aged subjects and ascribed it to dietary deficiency, but again the method employed was that of Ellinger and Dojmi (vide supra). The question of increased coproporphyrinuria in other diseases in which involvement of the skin is prominent has been reviewed by Dobriner (16i) who found conflicting reports concerning such conditions as erythema multiforme, lupus erythematosus, and xeroderma pigmentosum. Considerable variability was also evident in the rather extensive study of Brunsting and co-workers (10c). It may be noted with respect to this study that of nine cases of so called "eczema or urticaria solare", only two exhibited increases of porphyrin excretion and these were but slight. Goeckermann, Osterberg and Sheard (30) had previously noted a considerable increase in one such instance. Thus it appears that, as in hydroa aestivale seu vacciniforme, the porphyrins are implicated in some but not all cases of solar eczema or urticaria.

Origin and significance. It may be well in concluding this review to attempt to correlate what has gone before in terms of the more fundamental aspects of the coproporphyrins, i.e., their site of formation and functions. The evidence, however, is too scanty to permit conclusions and it is only possible to summarize the existing knowledge and theories.

It is clear that the coproporphyrins are easily synthesized from "niedrige Bausteine" by unicellular organisms, viz., yeast cells (24g, k, n), and various types of bacteria (11; 13; 42; 57). In this connection, it is of interest that yeast cells undergoing autolysis manufacture coproporphyrin I without reducing their formation of protoporphyrin 9 (type III) (24g). Bacteria, on the other hand, synthesize coproporphyrin III (42; 53a, b), at least insofar as is known at present. The significance of this difference has not been determined. Whether the synthesis of coproporphyrin is as easily achieved by animal cells is unknown. Recent studies have shown that protoporphyrin, at least, is readily formed from such simple precursors as acetates and glycine (4c; 81a). The transition to coproporphyrin is theoretically rather easy, requiring merely the addition of formic acid to the vinyl groups of protoporphyrin (24l). van den Bergh and his co-workers (96c) believed they were able to prove the occurrence of this transition in liver perfusion experiments, but their results could not be confirmed by Watson and co-workers (77; 100n).

Borst and Königsdörffer (7) observed coproporphyrin in the megaloblasts of embryonic bone marrow and of the bone marrow in the case Petry in whom pernicious anemia was said to have been important in the cause of death. The subsequent finding of an increased excretion of coproporphyrin I in pernicious anemia (16h, g, l; 98c; 100f), hemolytic anemias (16a, d, m; 100b), and regenerative post hemorrhagic anemias (16h), all of which are characterized by increased, albeit differing varieties of erythropoiesis in the bone marrow, led to postulation that formation of coproporphyrin I is proportional to the rate of

erythropoiesis (16d, h, m; 100f). Rimington (73e) and Dobriner and Rhoads (16g) advanced the theory that coproporphyrin I is formed as a by-product of hemoglobin synthesis, Rimington (73e) suggesting a normal ratio of 10,000 of type III (protoporphyrin) to 1 of type I (coproporphyrin). The constant proportion in rate of formation of these two porphyrins noted by Dobriner and Rhoads led them to suggest, quite similarly, that "the two substances are not synthesized independently but are products of the same chemical process". As shown in figure 5, they postulated that two dipyrrylmethenes, A and B, are the precursors of both types I and III porphyrins. The main synthesis proceeding to the protoporphyrin 9 of the hemoglobin molecule would thus depend upon coupling

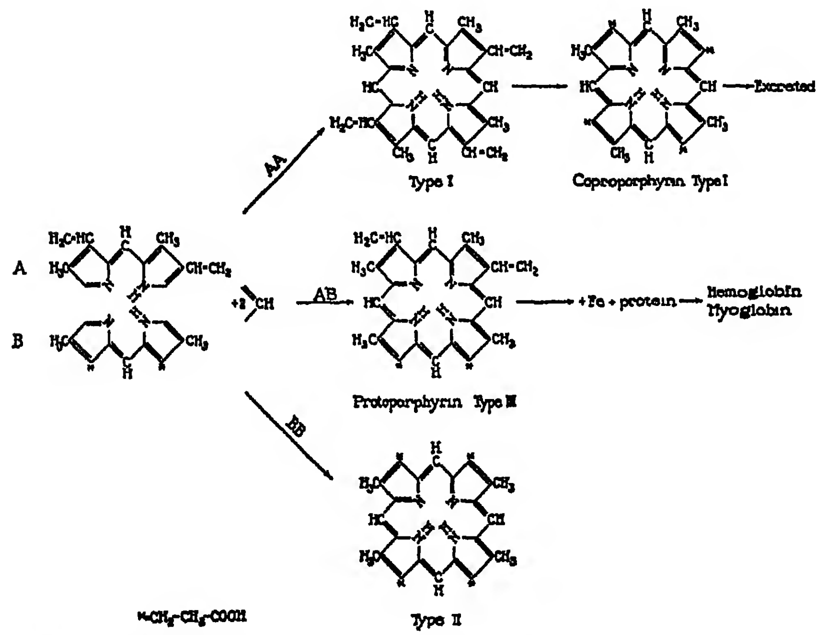


Fig. 5. Scheme of formation of hemoglobin and the coproporphyrins as according to Dobriner and Rhoads, J. Clin. Investigation 17: 95, 1935. Reproduced by permission of the publishers.

of two molecules A and B while the formation of coproporphyrin I would depend upon coupling of two molecules of A. It is further seen that two molecules of B would form a type II porphyrin. As already noted, however, there is no evidence for the occurrence of the latter type in nature. The scheme shown in figure 5 does not attempt to answer the question as to whether coproporphyrin I is formed by direct synthesis or by addition of formic acid to the vinyl groups of an isomeric (type I) protoporphyrin. According to Rimington's theory (73e) iron is "blocked" by certain substances, such as lead, from combination with protoporphyrin, and the latter, which is thus not used in hemoglobin synthesis, is converted to coproporphyrin III. It should be noted that this is regarded as a

disturbance in "Aufbau" rather than "Abbau". The latter, however, would be represented by the transition: hemoglobin → methemoglobin → coproporphyrin III, which Rimington and Hemmings (73c) and Brownlee (8b) believe to be the mode of elaboration induced by the sulfonamides and similar chemicals. Rimington's theory assumes that porphyrin formation is under an enzymic control, the nature of which is such that under normal circumstances, the above mentioned ratio of 10,000 to 1 prevails. According to this, in quite the same manner as postulated by Rhoads and Dobriner, an increased excretion of coproporphyrin I would be expected in conditions associated with increased erythropoiesis. Further studies of fecal and urinary coproporphyrin excretion are needed in a variety of anemias having both increased and decreased erythropoiesis. Preliminary studies (80g) have demonstrated that, in animals having diminished erythropoiesis due to radiation, there are significant decreases in coproporphyrin excretion as might be expected on the basis of these theories. This, however, has not been shown for human aplastic or hyporegenerative anemias.

Rimington's enzymic theory ascribes the marked increases of type I isomer in congenital porphyria to an inborn disturbance in enzymic control of the direction of synthesis. The over-production of type III in other cases of porphyria is viewed by Rimington as an abnormality in enzymic control of heme synthesis similar to that already referred to for lead poisoning, i.e., a "blocking" of the entrance of iron into variable fractions of the protoporphyrin molecule, which is then converted to coproporphyrin III. Actual evidence for this, however, is lacking, and it should be noted that Kench and co-workers (48) doubt the "blocking" theory because they failed to find any quantitative relationship in lead poisoning, between hemoglobin metabolism and coproporphyrin III excretion.

The fundamental position of uroporphyrin and its relation to coproporphyrin has not been determined. H. Fischer (24b) raised the question as to whether the latter was converted to the former by the kidneys, since at that time uroporphyrin had only been observed in the urine. Since it has been established that it is a normal constituent of mammalian embryonic and early postnatal bone and that it is still present in the ossicles of the inner ear in later postnatal life (51a), further, that in porphyria it is commonly encountered in the feces (80a; 100q) and in the viscera, especially the liver (7; 24q; 100q), the important question for future studies to answer is whether coproporphyrin formation proceeds over uroporphyrin by simple decarboxylation. While uroporphyrin probably has some function in pre- and early postnatal life, the possibility exists that under normal circumstances in later postnatal life, it is formed and rapidly decarboxylated to coproporphyrin. Porphyria, on this basis, would consist of a marked over production of uroporphyrin with varying degrees of failure of decarboxylation to coproporphyrin, in different cases.

The possibility also exists, however, that the essential constitutional error in porphyria is a further carboxylation of copro- to uroporphyrin. This would have to include, of course, an excessive production of coproporphyrin and a variable efficiency of carboxylation in order to explain cases of the type repre-

sented in figure 4. From a strictly chemical standpoint, the primary synthesis of uroporphyrin with secondary decarboxylation would appear to be the more likely of the two possibilities. This needs further study as a possible normal mechanism of coproporphyrin formation.

The possibility of formation of a fraction of the urinary coproporphyrin in the intestinal tract either from hemoglobin, or as a product of bacterial metabolism, at least under certain circumstances, cannot be wholly excluded. Yet it can be said that the contribution from either of these sources is at least quite small and that the urinary coproporphyrin is mainly endogenous. It is probable that coproporphyrin III is formed in small amount both from blood in the intestine and by certain of the intestinal flora, but the most recent objective studies (53a, b) have failed to reveal any appreciable absorption and urinary excretion. It may be noted, however, that the discrepancy between these results and those reported by Kaunitz (46), is unexplained. Mallinckrodt-Haupt (57) stated that the increased coproporphyrin (I) of the urine in pernicious anemia is derived entirely from an abnormal bacterial flora in the colon, the abnormality, in his opinion, depending upon the achlorhydria. He implied, in fact, that all of the urinary coproporphyrin, whether normally or in disease, was derived in this way. As already observed, the available evidence indicates that whatever coproporphyrin is formed by the bacterial flora is type III rather than type I, and further that it is very questionable whether this is absorbed. There are in addition, a number of other objections to this view, as follows:

1. The meconium, which does not have a bacterial flora is relatively rich in coproporphyrin I (79b; 99f), a finding which we have confirmed.[5] Garrod (27c) and Günther (34a) had previously observed "hematoporphyrin" in the meconium, and Schumm (79b) first identified it as coproporphyrin.

2. The excretion of coproporphyrin I in the feces and urine of pernicious anemia patients decreases significantly after liver induced remission (16e; 100d) although the achlorhydria and hence the effect on the bacterial flora of the colon which Mallinckrodt-Haupt emphasizes, is permanent.

3. The amounts of coproporphyrin I in the feces of hemolytic jaundice patients usually exceed considerably those of pernicious anemia in relapse (100d, e, f), yet the gastric hydrochloric acid and the bacterial flora of the colon in this disease are, usually at least, quite normal.

4. If the bacterial flora of the intestine were mainly responsible for the urinary coproporphyrin, one would have to assume a differing flora in such conditions as aplastic or refractory anemia on the one hand from that in infectious hepatitis on the other, or in infectious hepatitis from that in poliomyelitis, in accordance with the recent findings as referred to in the foregoing. While such a difference is barely possible, it would appear to be quite unlikely.

With the possible exception of a derivation of coproporphyrin III from methemoglobin, as discussed previously, there is little evidence for, and a great deal against the catabolic theory of coproporphyrin formation. It appears that

[5] Unpublished.

Schreus (78b) was the last to champion this belief, although without offering any definite support for it. The earlier investigators, viz., Stokvis, Garrod, Saillet, Günther and others had rather naturally assumed such an origin since, as previously pointed out, they had all confused copro- with hematoporphyrin, and the derivation of the latter from hemoglobin, in vitro, was one of the few things about the porphyrins that was clearly appreciated in that period. Schreus (78b, c, g), and Carrié (11), although well aware of the nature of the coproporphyrins, believed that they were derived from hemoglobin destruction principally because of the speed of increase in the urine after onset of the disease process, and because of their data indicating a proportionality between rate of blood destruction and amount of coproporphyrin excreted. Insofar as the type I isomer is concerned, it is clear that this could be related just as well to the rate of erythropoiesis, as emphasized by the theories of Rimington and Dobriner (vide supra). Furthermore, the data of others (48; 100e) have not revealed correlation between the rate of hemoglobin catabolism and the excretion of coproporphyrin III either in urine or feces.

Klüver's discovery (51a, b) of the regular occurrence of small amounts of coproporphyrin in the central nervous system of warm blooded animals, and the tentative identification as the type III isomer together with the increased excretion of this isomer in poliomyelitis, lead one to question whether the urinary porphyrin may be derived at least in part from that of the nervous system. The different orders of magnitude of the concentrations in the brain and the urine speak against this possibility, but perhaps do not exclude it. Nothing is known of the rate of formation of coproporphyrin, and if Granick and Gilder's hypothesis (31) is correct, that coproporphyrin may serve as a regulator of oxygen consumption by living cells, it would be conceivable that nervous tissue might produce and liberate it much more rapidly under certain circumstances. As yet, nothing is known as to possible alterations of central nervous system coproporphyrin concentration in disease.

Prior to Klüver's discovery, Keilin (47) had suggested that cytochrome might be the parent substance of coproporphyrin. Klüver (51b) pointed out, however, that the areas relatively rich in cytochrome were poor in coproporphyrin, or devoid of it entirely, and vice versa.

Reasons have been given in the foregoing for ascribing the increased urinary coproporphyrin I in hepatitis and many cases of cirrhosis, to hepato-cellular functional impairment. It is impressive, however, that a chronic alcoholic with advanced cirrhosis and obvious liver functional impairment, usually does not excrete increased type I, but rather a considerable amount of the type III isomer. One may contrast two cases of cirrhosis exhibiting about equally severe disturbance of hepatic function in other respects, one, a chronic alcoholic excreting 150 γ of type III and but 50 of type I, the other, a non-alcoholic whose cirrhosis followed an infectious hepatitis, excreting 175 γ of type I and 15 γ of type III. This suggests that various disease agents drive or direct the coproporphyrin formation one way or the other, and that there may be a closer relationship between the formation of the two isomers than hitherto considered.

Further studies should seek to determine whether there is a balanced production of the coproporphyrins normally favoring type I, and whether disturbances in this ratio are subject in part to constitutional factors. The latter appear requisite to explain the striking differences thus far encountered in such well defined entities as Hodgkin's disease and hemochromatosis. At the same time it is recognized that other and as yet unknown factors may be responsible.

Thus the exact site and mode of formation of the coproporphyrins and their physiological rôle, remain in doubt. It would appear that microfluorospectrometry and the use of synthetic porphyrins containing C_{14}, are of most promise for the immediate future investigation of these questions.

REFERENCES

(1) Barber, H. W. and Z. Howitt. Guy's Hosp. Rep. **76:** 314, 1926.
(2) Bassi, U. Clin. Med. et al. **65:** 241, 1934.
(3) Beckh, W., P. Ellinger and T. D. Spies. Quart. J. Med. **30:** 305, 1937.
(4a) Bingel, A. Fortschr. d. Neurol. **9:** 265, 1937.
(b) Bingel, A. Ztschr. f. d. ges. Neurol. u. Psychiat. **158:** 79, 1937.
(c) Block, K. E. and D. Rittenberg. J. Biol. Chem. **159:** 45, 1945.
(5a) Blum, H. F. Photodynamic action and diseases caused by light. New York, Reinhold Publishing Corp., 1941.
(b) Blum, H. F. and H. J. Templeton. J. A. M. A. **108:** 548, 1937.
(6) Boas, I. Deutsch. med. Wchnschr. **4:** 126, 1933.
(7) Borst, M. and H. Königsdorffer. Untersuchungen über Porphyrie. S. Hirzel, Leipzig, 1929.
(8a) Brownlee, G. Proc. Roy. Soc. Med. **32:** 1276, 1939.
(b) Brownlee, G. Biochem. J. **33:** 697, 1939.
(9a) Brugsch, J. T. Ztschr. f. d. ges. exper. Med. **95:** 471, 1935.
(b) Brugsch, J. T. Ztschr. f. d. ges. exper. Med. **95:** 482, 1935.
(c) Brugsch, J. T. Ergebn. d. ges. Med. **20:** 424, 1935.
(d) Brugsch, J. T. Ergebn. d. inn. Med. u. Kinderh. **51:** 86, 1936.
(e) Brugsch, J. T. Ztschr. f. physiol. Chem. **99:** 585, 1936.
(f) Brugsch, J. T. and A. Keys. Proc. Staff. Meetings Mayo Clinics **12:** 609, 1937.
(g) Brugsch, J. T. Mensch u. Chlorophyll; Ergeb. d. inn. Med. u. Kinderheilk. **56:** 614, 1939.
(10a) Brunsting, L. A. Proc. Staff Meetings Mayo Clinic **12:** 614, 1937.
(b) Brunsting, L. A. and H. L. Mason. J. A. M. A. **132:** 509, 1946.
(c) Brunsting, L. A., J. T. Brugsch and P. A. O'Leary. Arch. Dermatol. and Syph. **39:** 294, 1939.
(11) Carrié, C. Die Porphyrine. G. Thieme, Liepzig, 1936.
(11a) Chandler, F. G., C. A. Harrison and C. Rimington. Brit. Med. J. **2:** 1173, 1939.
(b) Cartwright, G. E., M. A. Lauritsen, P. J. Jones, I. M. Merrill and M. M. Wintrobe. J. Clin. Investigation **25:** 65, 1946.
(12) Chu, E. J. H. and C. J. Watson. To be published.
(13) Coulter, C. B. and F. M. Stone. Proc. Soc. Exper. Biol. and Med. **38:** 423, 1938.
(14) Crawford, H. Med. J. Australia **1:** 580, 1933.
(15a) Derrien, E. and C. Benoit. Bull. Soc. Chim. Fr. **43:** 522, 1928.
(b) Derrien, E. and P. Cristol. C. R. Soc. Biol. Paris **103:** 126, 1930.
(16a) Dobriner, K. J. Biol. Chem. **113:** 1, 1936.
(b) Dobriner, K. Proc. Soc. Exper. Biol. and Med. **35:** 175, 1936.
(c) Dobriner, K. J. Biol. Chem. **120:** 115, 1937.
(d) Dobriner, K. Proc. Soc. Exper. Biol. and Med. **36:** 757, 1937.
(e) Dobriner, K. and W. H. Barker. Proc. Soc. Exper. Biol. and Med. **36:** 864, 1937.

(f) DOBRINER, K. AND C. P. RHOADS. New England Med. J. 219: 1027, 1938.
(g) DOBRINER, K. AND C. P. RHOADS. J. Clin. Investigation 17: 95, 1938.
DOBRINER, K. AND C. P. RHOADS. J. Clin. Investigation 17: 105, 1938.
DOBRINER, K. AND C. P. RHOADS. Physiol. Rev. 20: 416, 1940.
DOBRINER, K., C. P. RHOADS AND L. E. HUMMEL. J. Clin. Investigation 17: 125, 1938.
(k) DOBRINER, K., W. H. STRAIN AND S. A. LOCALIO. Proc. Soc. Exper. Biol. and Med. 38: 748, 1938.
(l) DOBRINER, K., W. H. STRAIN, H. GUILD AND S. A. LOCALIO. J. Clin. Investigation 17: 761, 1938.
(m) DOBRINER, K., W. H. STRAIN, S. A. LOCALIO, H. KEUTMANN AND D. I. STEPHENS. Proc. Soc. Exper. Biol. and Med. 36: 755, 1937.
(n) DOBRINER, K., W. H. STRAIN, AND S. A. LOCALIO. Proc. Soc. Exper. Biol. and Med. 36: 752, 1937.
(17a) DUESBERG, R. Arch. f. exper. Path. u. Pharmakol. 162: 249, 1931.
(b) DUESBERG, R. München med. Wchnschr. 79: 1821, 1932.
(18) EICHLER, P. Ztschr. Neurol. 141: 363, 1932.
(19) ELLINGER, A. AND O. RIESSER. Ztschr. f. physiol. Chem. 98: 1, 1916.
(20) ELLINGER, P. AND F. DOJMI. Chem. and Industry 13: 507, 1935.
(21) FIGGE, F. H. G., T. H. CAREY AND G. S. WEILAND. J. Lab. and Clin. Med. 31: 752, 1946.
(22) FIKENTSCHER, R. Biochem. Ztschr. 249: 257, 1932.
(23) FINK, H. AND W. HOERBURGER. Die Naturwissenschaften 18: 292, 1934.
(24a) FISCHER, H. Ztschr. f. physiol. Chem. 95: 34, 1915.
(b) FISCHER, H. Ztschr. f. physiol. Chem. 96: 148, 1915.
(c) FISCHER, H. Ztschr. f. physiol. Chem. 97: 109, 1916.
(d) FISCHER, H. Ztschr. f. physiol. Chem. 155: 96, 1926.
(e) FISCHER, H. Oppenheimer's Handb. d. Biochem. des Menschen und die Tiere. Zweite Auflage, Ergänzungsband p. 72, G. Fischer, Jena, 1930.
(f) FISCHER, H. AND E. DUESBERG. Arch. f. exper. Path. u. Pharmakol. 166: 95, 1932.
(g) FISCHER, H. AND H. FINK. Ztschr. f. physiol. Chem. 150: 243, 1925.
(h) FISCHER, H. AND H. HILMER. Ztschr. f. physiol. Chem. 153: 167, 1926.
(i) FISCHER, H. AND J. HIERNEIS. Ztschr. f. physiol. Chem. 196: 155, 1931.
(j) FISCHER, H. AND LIBOWITZKY. Ztschr. f. physiol. Chem. 241: 220, 1936.
(k) FISCHER, H. AND H. ORTH. Die Chemie des Pyrrols. Bd II; I Hälfte, Leipzig: Akad. Verlagsgesellsch., 1937.
(l) FISCHER, H., K. PLATZ AND K. MORGENROTH. Ztschr. f. physiol. Chem. 182: 265, 1929.
(m) FISCHER, H. AND W. ZERWECK. Ztschr. f. physiol. Chem. 132: 12, 1924.
(n) FISCHER, H. AND K. SCHNELLER. Ztschr. f. physiol. Chem. 135: 253, 1924.
(o) FISCHER, H. AND W. ZERWECK. Ztschr. f. physiol. Chem. 137: 176, 1924.
(p) FISCHER, H. AND K. ZEILE. Liebig's Ann. d. Chem. 468: 98, 1929.
(q) FISCHER, H., H. HILMER, F. LINDNER AND B. PÜTZER. Ztschr. f. physiol. Chem. 150: 44, 1925.
(25a) FRANKE, K. Ztschr. f. klin. Med. 130: 222, 1936.
(b) FRANKE, K. AND R. FIKENTSCHER. München med. Wchnschr. 82: 171, 1935.
(c) FRANKE, K. AND S. LITZNER. Ztschr. f. klin. Med. 129: 115, 1935.
(26) GAFFRON, H. Ztschr. f. physiol. Chem. 179: 157, 1926.
(27a) GARROD, A. E. J. Physiol. 13: 598, 1892.
(b) GARROD, A. E. J. Path. and Bact. 1: 87, 1893.
(c) GARROD, A. E Lancet 2: 1323, 1900.
(d) GARROD, A. E. Inborn errors of metabolism. H. Froude, London 1923. 2nd ed.
(28) GILDEMEISTER, H. Ztschr. f. d. ges. exper. Med. 102: 58, 1937.
(29) GILLAM, A. E. AND J. E. KENCH. Lancet, London 2: 806, 1940.

(30) Goeckerman, H. W., A. E. Osterberg and C. Sheard. Arch. Dermat. and Syph. 20: 501, 1929.
(31) Granick, S. and H. Gilder. Science 101: 540, 1945.
(32a) Grinstein, M. La Prensa Medica 44: 2241, 1940.
(b) Grinstein, M. La Prensa Medica 45: 2295, 1940.
(c) Grinstein, M., S. Schwartz and C. J. Watson. J. Biol. Chem. 157: 323, 1945.
(d) Grinstein, M. Revista Arg.—Norteamic. de Cienc Med. 9: 2, 1944.
(33a) Grotepass, W. Ztschr. f. physiol. Chem. 205: 193, 1932.
(b) Grotepass, W. Ztschr. f. physiol. Chem. 253: 276, 1938.
(c) Grotepass, W. and A. Defalque. Ztschr. f. physiol. Chem. 252: 155, 1938.
(34a) Günther, H. Ergebn. allg. Path. u. path. Anat. 20: 608, 1922.
(b) Günther, H. Krankheiten des Blutes und der blutbildenden Organe. Bd 2, Berlin: Julius Springer, 1925.
(35) Halter, K. Arch. f. Dermatol. u. Syph. 178: 340, 1938.
(36) Haranghy, L. Centralbl. f. allg. Path. u. path. Anat. 54: 161, 1932.
(37a) Hausmann, W. Wien. klin. Wchnschr. 22: 1820, 1909.
(b) Hausmann, W. Sonderbande z. Strahlentherapie. Wien-Berlin: Urban and Schwarzenberg, 1923.
(c) Hausmann, W. and H. Haxthausen. Die Lichtkrankungen der Haut. Urban and Schwarzenberg; Berlin and Vienna, 1929.
(d) Hausmann, W. and F. M. Kuen. Biochem. Ztschr. 265: 105, 1933.
(38) Haxthausen, H. Dermat. Wchnschr. 84: 827, 1927.
(38a) Hill, R. and H. F. Holden. Biochem. J. 20: 1326, 1926.
(39a) Hoerburger, W. Inaugural Dissertation, Erlangen, 1933.
(b) Hoerburger, W. and H. Fink. Ztschr. f. physiol. Chem. 236: 136, 1935.
(40) Hoesch, K. and C. Carrié. Ztschr. f. klin. Med. 129: 214, 1935.
(41) Hoppe-Seyler, F. Cited by Nencki and Sieber (64).
(42) Jakob, A. Klin. Wchnschr. 18: 1024, 1939.
(43a) Kämmerer, H. Klin Wchnschr. 2: 1153, 1923.
(b) Kämmerer, H. and W. K. Meyer. Deutsch. Arch. f. klin. Med. 179: 392, 1936.
(44a) Kapp, E. M. and A. F. Coburn. Brit. J. Exper. Path. 17: 255, 1936.
(b) Kapp, E. M. Brit. J. Exper. Path. 20: 33, 1939.
(45a) Kark, R. and A. P. Meiklejohn. Am. J. Med. Sc. 201: 380, 1941.
(b) Kark, R. and A. P. Meiklejohn. J. Clin. Investigation 21: 91, 1942.
(46) Kaunitz, H. Ztschr. f. klin. Med. 133: 552, 1938.
(47) Keilin, D. Proc. Roy. Soc. London (B) 98: 312, 1925.
(48) Kench, J. E., A. E. Gillam and R. E. Lane. Biochem. J. 36: 384, 1942.
(49) Keys, A. and J. Brugsch. J. Am. Chem. Soc. 60: 2135, 1938.
(50) Kirstahler, A. Tab. Biologicae Periodicae 7: 48, 1931.
(51a) Klüver, H. J. Psychol. 17: 209, 1944.
(b) Klüver, H. Science 99: 482, 1944.
(c) Klüver, H. Personal communication.
(52) Lageder, K. Arch. f. Verdauungskr. 56: 237, 1934.
(53a) Larson, E. A. Graduate Thesis—Univ. of Minn., November 1946.
(b) Larson, E. A. and C. J. Watson. Studies of Coproporphyrin (V) (to be published).
(54) Laubender, W. and K. Monden. Arch. f. exper. Path. u. Pharmakol. 188: 562, 1938.
(55) Libowitzky, H. and K. F. Scheid. Klin. Wchnschr. 17: 156, 1938.
(56) MacMunn, C. A. Proc. Roy. Soc. of London 31: 206, 1880.
(57) v. Mallinckrodt-Haupt, A. S. Klin. Wchnschr. 20: 190, 1941.
(58a) Mason, H. L. and S. Nesbitt. J. Biol. Chem. 152: 19, 1944.
(b) Mason, V. R., C. Courville and E. Ziskind. Medicine 12: 355, 1933.
(59) Massa, M. Riforma med. 48: 1669, 1932.

(60a) MERTENS, E. Klin. Wchnschr. **16**: 61, 1936.
(b) MERTENS, E. Ztschr. f. physiol. Chem. **250**: 57, 1937.
(61) MEYER-BETZ, F. Deutsch. Arch. f. klin. Med. **112**: 476, 1913.
(62) MICHELI, F. AND G. DOMINICI. Deutsch. Arch. f. klin. Med. **171**: 154, 1931.
(63) MINIBECK, H. Biol. Ztschr. **293**: 219, 1937.
(64) NENCKI, M. AND N. SIEBER. Arch. f. exp. Path. u. Pharmakol. **24**: 430, 1888.
(65) NESBITT, S. AND A. M. SNELL. Arch. Int. Med. **69**: 573, 1942.
(66) NEUBAUER, O. Arch. f. exper. Path. u. Pharmakol. **43**: 456, 1900.
(67) OTTO, H. Arch. f. Gewerbepath. u. Hyg. **8**: 655, 1938.
(68) PENEW, L. AND C. TROPP. Deutsch. Arch. f. klin. Med. **180**: 423, 1937.
(69) POULIKAKOS, P. AND C. TROPP. Deutsch. Arch. f. klin. Med. **183**: 342, 1938.
(70) RAFSKY, H. A. AND B. NEWMAN. Am. J. Digest. Dis. **9**: 43, 1942.
(71) RASK, E. N. AND W. H. HOWELL. Am. J. Physiol. **84**: 363, 1928.
(72) REITLINGER, K. AND P. KLEE. Arch. f. exper. Path. u. Pharmakol. **127**: 277, 1938.
(73a) RIMINGTON, C. Onderstepoort J. Vet. Sc. **7**: 567, 1936.
(b) RIMINGTON, C. Ann. Rev. Biochem. **12**: 425, 1943.
(c) RIMINGTON, C. AND A. W. HEMMINGS. Lancet **1**: 770, 1938.
(d) RIMINGTON, C. AND A. W. HEMMINGS. Biochem. J. **33**: 960, 1939.
(e) RIMINGTON, C. C. r. Lab. Carlsberg, Sér. chim. **22**: 454, 1938.
(74) ROSENBLUM, L. A. AND N. JOLLIFFE. Am. J. Med. Sc. **199**: 853, 1940.
(75a) ROTH, E. Deutsch. Arch. f. klin. Med. **178**: 185, 1935.
(b) ROTH, E. Ztschr. f. klin. Med. **129**: 123, 1936.
(76) SAILLET. Rev. de. méd. **16**: 542, 1896.
(77) SALZBURG, P. AND C. J. WATSON. J. Biol. Chem. **139**: 593, 1941.
(78a) SCHREUS, H. T. Klin. Wchnschr. **4**: 121, 1934.
(b) SCHREUS, H. T. Klin. Wchnschr. **13**: 121, 1934
(c) SCHREUS, H. T. Klin. Wchnschr. **14**: 1717, 1935.
(d) SCHREUS, H. T. AND C. CARRIÉ. Dermat. Ztschr. **62**: 357, 1931.
(e) SCHREUS, H. T. AND C. CARRIÉ. Ztschr. f. klin. Med. **125**: 330, 1935.
(f) SCHREUS, H. T. AND C. CARRIÉ. Klin. Wchnschr. **11**: 1017, 1931.
(g) SCHREUS, H. T. Klin. Wchnschr. **13**: 334, 1934.
(79a) SCHUMM, O. Ztschr. f. physiol. Chem. **105**: 168, 1919.
(b) SCHUMM, O. Ztschr. f. physiol. Chem. **126**: 109, 1923.
(c) SCHUMM, O. Arch. f. exper. Path. u. Pharmakol. **191**: 529, 1939.
(d) SCHUMM, O. Ztschr. f. d. ges. exper. Med. **106**: 252, 1939.
(80a) SCHWARTZ, S. AND C. J. WATSON. Proc. Soc. Exper. Biol. and Med. **47**: 390, 1941.
(b) SCHWARTZ, S., V. E. HAWKINSON AND C. J. WATSON. Science **103**: 338, 1946.
(c) SCHWARTZ, S. AND R. ZAGARIA. Plutonium Project Record Vol. 22, B (to be published).
(d) SCHWARTZ, S. Plut. Proj. Record, Vol. 20 (to be published).
(e) SCHWARTZ, S., V. E. HAWKINSON, S. COHEN AND C. J. WATSON. J. Biol. Chem. **168**: 133, 1947.
(f) SCHWARTZ, S., R. HUNTER, M. GLICKMAN AND J. WALLACE. Plut. Proj. Record Vol. 22 (to be published).
(g) SCHWARTZ, S., R. ZAGARIA AND C. J. WATSON. Plutonium Proj. R., v. 22 (to be published).
(81) SCOLARI, E. Giorn. Ital. Dermatol. Sifilog. **78**: 5, 1937.
(81a) SHEMIN, D. AND D. RITTENBERG. J. Biol. Chem. **159**: 567, 1945.
(82) SHIBUYA, H. Strahlentherapie **17**: 412, 1924.
(83) SIEDEL, W. Die Chemie **56**: 169, 1943.
(84) SILVER, B. AND M. ELLIOTT. J. A. M. A. **112**: 723, 1939.
(85) SPIES, T. D., Y. SASAKI AND E. GROSS. South. Med. J. **31**: 483, 1938.
(86a) STOKVIS, B. J. Ztschr. f. klin. Med. **28**: 1, 1895.
(b) STOKVIS, B. J. Ned. Tijdschr. Geneesk. **13**: 413, 1889.

(87a) Stokes, J. H., H. Beerman and N. R. Ingraham. Am. J. Med. Sc. **203**: 608, 1942.
(b) Stokes, J. H., H. Beerman and N. R. Ingraham. Am. J. Med. Sc. **204**: 601, 1942.
(88) Strauch, C. B. Am. J. Dis. Child. **40**: 800, 1930.
(89) Sumegi, S. and J. Putnoky. Arch. f. Gewerbepath. u. Hyg. **9**: 566, 1939.
(90) Sutherland, D., V. E. Hawkinson and C. J. Watson. Studies of Coproporphyrin (IV) (to be published).
(91) Taylor, J. J., M. L. Solomon, G. S. Wieland and F. J. Figge. J.A.M.A. **131**: 26, 1946.
(92a) Thiel, W. Verh. deutsch. Ges. inn. Med. 81, 1933.
(b) Thiel, W. Klin. Wchnschr. **19**: 1, 1934.
(93a) Tropp, C. and K. Siegler. Deutsch. Arch. f. klin Med. **180**: 402, 1937.
(b) Tropp, C. and L. Penew. Deutsch. Arch. f. klin Med. **180**: 411, 1937.
(94) Turner, W. J. J. Lab. and Clin. Med. **26**: 323, 1940.
(95a) Urbach, E. Klin. Wchnschr. **17**: 304, 1938.
(b) Urbach, E. and J. Blöch. Wien. klin. Wchnschr. **47**: 527, 1934.
(96a) van den Bergh, H. and W. Grotepass. Klin. Wchnschr. **15**: 586, 1933.
(b) van den Bergh, H. and M. Hyman. Deutsch. med. Wchnschr. **36**: 1492, 1928.
(c) van den Bergh, H., W. Grotepass and F. E. Revers. Klin. Wchnschr. **11**: 1534, 1932.
(d) van den Bergh, H., R. Regniers and W. Muller. Arch. f. Verdauungskr. **42**: 302, 1928.
(97a) Vannotti, A. Ergebn. d. inn. Med. u. Kinderh. **49**: 337, 1935.
(b) Vannotti, A. Arch f. gewerbepath. u. Hyg. **8**: 266, 1937.
(c) Vannotti, A. Porphyrine und Porphyrinkrankheiten. J. Springer, Berlin, 1937.
(98a) Vigliani, E. C. Arch. per le Sc. Med. **55**: 391, 1938.
(b) Vigliani, E. C. and C. Angeleri. Klin. Wchnschr. **15**: 700, 1936.
(c) Vigliani, E. C. and H. Libowitsky. Klin. Wchnschr. **16**: 1243, 1937.
(d) Vigliani, E. C. and B. Sonzini. Arch. per le Sc. Med. **65**: 363, 1938.
(e) Vigliani, E. C. and J. Waldenström. Deutsch. Arch. f. klin. Med. **180**: 182, 1937.
(99a) Waldenström, J. Acta Med. Scandinav. **83**: 281, 1934.
(b) Waldenström, J. Deutsch. Arch. f. klin Med. **178**: 38, 1935.
(c) Waldenström, J. Acta Med. Scandinav. 1937, Supp. 82.
(d) Waldenström, J. and S. Wendt. Ztschr. f. physiol. Chem. **259**: 157, 1939.
(e) Waldenström, J., H. Fink and W. Hoerburger. Ztschr. f. physiol. Chem. **233**: 1, 1935.
(f) Waldenström, J. Ztschr. f. physiol. Chem. **239**: Supp. III, 1936.
(g) Waldenström, J. and B. Vahlquist. Ztschr. f. physiol. Chem. **260**: 189, 1939.
(100 a) Watson, C. J. Arch. Int. Med. **47**: 698, 1931.
(b) Watson, C. J. J. Clin. Investigation **14**: 106, 1935.
(c) Watson, C. J. J. Clin. Investigation **14**: 110, 1935.
(d) Watson, C. J. J. Clin. Investigation **14**: 116, 1935.
(e) Watson, C. J. J. Clin. Investigation **15**: 327, 1936.
(f) Watson, C. J. J. Clin. Investigation **16**: 383, 1937.
(g) Watson, C. J. Downey's Handbook of hematology, Vol. IV. Paul Hoeber, New York, 1938.
(h) Watson, C. J. Oxford Medicine **4**: 228, 1938.
(i) Watson, C. J. Proc. Soc. Exper. Biol. and Med. **39**: 514, 1938.
(j) Watson, C. J. Proc. Soc. Exper. Biol. and Med. **41**: 591, 1939.
(k) Watson, C. J. South. Med. J. **36**: 1, 1943.
(l) Watson, C. J. and J. A. Layne. Ann. Int. Med. **19**: 183, 1943.
(m) Watson, C. J. and S. Schwartz. Proc. Soc. Exper. Biol. and Med. **44**: 7, 1940.
(n) Watson, C. J., L. J. Pass and S. Schwartz. J. Biol. Chem. **139**: 583, 1941.
(o) Watson, C. J. and S. Schwartz. Abstract. Proc. Soc. Clin. Investigation, May, 1941.

(p) WATSON, C. J., M. GRINSTEIN AND V. E. HAWKINSON. J. Clin. Investigation 23: 69, 1944.
(q) WATSON, C. J., S. SCHWARTZ AND V. E. HAWKINSON. J. Biol. Chem. 157: 345, 1945.
(r) WATSON, C. J., W. SCHULZE, V. E. HAWKINSON AND A. B. BAKER. Proc. Soc. Exper. Biol. and Med. 64: 73, 1947.
(s) WATSON, C. J., S. SCHWARTZ, D. SUTHERLAND AND V. E. HAWKINSON. Studies of Coproporphyrin (I) (to be published).
(t) WATSON, C. J., R. C. CAPPS, E. M. RAPPAPORT AND V. E. HAWKINSON. Studies of Coproporphyrin (II) (to be published).
(u) WATSON, C. J., D. SUTHERLAND AND V. E. HAWKINSON. Studies of Coproporphyrin (III) (to be published).
(v) WATSON, C. J. AND V. E. HAWKINSON. Studies of Coproporphyrin (VI) (to be published).
(w) WATSON, C. J., R. ZAGARIA, B. MORGAN AND S. SCHWARTZ. Studies of Coproporphyrin (VII) (to be published).
(x) WATSON, C. J. AND V. E. HAWKINSON. Studies of Coproporphyrin (VIII) (to be published).
(y) WATSON, C. J. AND S. SCHWARTZ. Proc. Soc. Exper. Biol. and Med. 47: 393, 1941.
(101) WEISS, H. Deutsch. Arch. f. klin. Med. 149: 255, 1925.

PHYSIOLOGICAL REVIEWS

VOL. 27 OCTOBER, 1947 No. 4

BODY SIZE AND METABOLIC RATE

MAX KLEIBER

Division of Animal Husbandry, College of Agriculture, University of California, Davis

CORRELATION BETWEEN BODY SIZE AND METABOLIC RATE. Günther (1944) introduces a recent review on body weight and metabolic rate with a motto which starts as follows:

"It is believed that far greater progress will be made by discarding all thoughts of a uniformity in heat loss and emphasizing the non-uniformity in heat production. . . ."

The sentence is a citation from Benedict's book, *Vital Energetics* (1938, p. 194). It is rather difficult to understand how forgetting all thoughts of uniformity and emphasizing non-uniformity can stimulate a comparison of metabolic rates of large and small animals. Any comparison presupposes a common basis, and if I were convinced of the "futility of attempts to discover a unifying principle in metabolism" (Benedict, l.c., p. 178) I should not attempt to write a review on the relation of body size and metabolic rate.

The reader can be expected to spend time on this review only when he can be convinced that body size and metabolic rate are actually related. That these two variables are related is in fact common knowledge.

Does a horse produce more heat per day than a rat or do some rats produce more heat than do some horses? Almost anybody who understands what is meant by "heat production per day" will not hesitate to give the correct answer and will even be convinced that the daily rate of heat production of men or sheep is greater than that of rats, but smaller than that of horses. Thus most people (among those who understand the question) are convinced that in general the bigger homeotherms produce more heat per day than the smaller homeotherms, that, in other words, the metabolic rate of homeotherms is positively correlated to body size.

The answer to the next question: "does a horse produce more heat per day per kilogram of body weight than a rat?" requires some biological training. Most biologists, however, will not hesitate to answer that the rate of heat production per unit body weight of the big animal is less than that of the small animal.

The positive correlation between metabolic rate and body size, and the negative correlation between metabolic rate per unit weight and body size, establish two limits between which we expect to find the rate of heat production of a horse if we know the rate of heat production of a rat. We expect the metabolic rate of the horse to be somewhat between that of the rat, and that of the rat times the ratio of horse weight to rat weight, provided of course that we do not regard these two correlations as simply accidental.

If we are firmly convinced that the metabolic rate of horses, and other homeotherms of similar size, is never outside these two limits, then we admit to recognize a natural law between body size and metabolic rate. The firmness of the

conviction, or the probability of being correct, rather than the range within which we predict a result, justifies the term "law."

For the usefulness of a law, however, the accuracy of prediction is important. We may attempt to formulate the relation between body size and metabolic rate more precisely than "somewhere in the range between rate per animal and rate per unit weight."

Encouraging for such attempts is the fact that the correlation between metabolic rate and body weight is, indeed, rather high. For 26 groups of mammals, dealt with particularly in this review, the coefficient of correlation between the two variables amounts to +0.98.

This high correlation involves standardization of conditions under which metabolic rates are measured. The standard metabolic rate should be measured while the animal is in a post absorptive condition, and kept within the range of metabolically indifferent environmental temperature.

The task is to find a metabolic body size which is chosen so that the metabolic rate per unit of this body size is the same for large and small animals.

The square meter of body surface is such a unit which allows a more accurate prediction of metabolic rate than the two limits mentioned above. The surface area of animals is, however, ill defined, and relatively recent results indicate that the metabolic rate per unit surface area tends to be greater the larger the animals (Kleiber, 1932). Among the 26 groups of animals chosen for discussion in this review the metabolic rate per unit of the 2/3 power of body weight (which roughly represents surface area) has a highly significant correlation to body weight. The correlation coefficient amounts to +0.71.

Such correlation disappears when the metabolic rate is divided by the 3/4 power of body weight. The unit of the 3/4 power of body weight, $kg.^{3/4}$, is therefore a suitable unit of metabolic body size.

Fasting homeotherms under standard conditions produce daily an average of about 70 kcal. of heat per $kg.^{3/4}$ or 3 kcal. per $kg.^{3/4}$ per hour.

The unit of metabolic body size is of great interest in comparative physiology. It permits an expression of the metabolic level of an animal independent of its body size, and it makes possible the comparison of a particular metabolic level with the interspecific mean of 70 kcal. per $kg.^{3/4}$ per day.

Such a comparison, furthermore, permits for example the statement that the metabolic rate of a 70 ton whale estimated by Irving, Scholander and Grinnell (1941) is high in comparison to that of other mammals (whereas the whale's rate per kg. is only 1/10 that of a mouse). Such a comparison, made possible by the use of metabolic body size, may stimulate research to establish the conditions producing that high metabolic rate. It may lead, for example, to a comparison with the metabolic rate of Eskimos which also is at a relatively high level (Rabinowitch and Smith, 1936). The unit of metabolic body size is further useful in evaluating levels of food intake in animal production, and in classifying farm animals with regard to their efficiency as food utilizers. Food requirements and dosages of most vitamins and drugs may be expressed in terms of metabolic body size.

This review is limited to mammals. The conclusions are, however, also applicable to birds whose metabolic level was found to be in line with the mammalian metabolic level (Kleiber, 1932). More recent results confirm the metabolic similarity of these two groups of homeotherms. From data published by Dukes (1937) one may calculate that the metabolic rate of hens is 73 kcal. per kg.$^{3/4}$ A rule of body size and metabolic rate, very similar to that for mammals and birds, has also been noted in heterotherms such as the crab by Weymouth and his co-workers (1944).

THE DISCOVERY OF THE SURFACE LAW. The metabolic rate (heat production per unit of time), in particular the basal metabolic rate of humans, is generally expressed in kilocalories per square meter of body surface. This procedure is based on the theory that in animals of different body size the metabolic rate is proportional to their respective surface areas. This theory, called the surface law, is now a little over a century old. It has its roots in the time when a vigorous belief in the power of reason inspired men to explain nature's aims and means. A good deal of a new theology was then evident in teleological scientific discussions. A theology in which Nature, as a bright goddess, revealed her aims and ways to a scientist as a more stern and sombre Lord had revealed His aims and wishes to the nonscientific or antiscientific priests.

Sarrus, a professor of mathematics, and Rameaux, a doctor of medicine and of science, both in Strasbourg, had Thillaye read a thesis to the Royal Academy of France during its meeting of July 23, 1839.

The authors apparently take it for granted that it is nature's aim to make the rate of heat production of large and small animals in proportion to their respective surface areas or the 2/3 power of their body weights. They then calculate how nature achieves this aim by modulating the frequency of heart beat and the stroke volume, always choosing the middle way between two possibilities. Their basic theory is worth being quoted: (p. 1098) "Lorsque la nature peut attaindre un but par plusieurs moyens, elle n'emploie jamais exclusivement l'un d'eux jusqu'aux limites, elle les fait concourir de manière que chacun de ces moyens tend à produire une part égale de l'effet total."

(When nature can achieve an aim by various means she never uses one of these means exclusively to the limit, she makes these means compete so that each one of them produces an equal part of the total effect.)

Bergmann and Leuckart (1855) concluded from measurements of Regnault and Reiset that the metabolic rate per unit weight was especially great for small animals. In four days sparrows consumed as much oxygen as they weighed. This today would be considered a very high metabolic rate, about four times as high as the rate observed in a sparrow by Benedict (1938). Richet (1889) discovered "après coup," that is empirically, that the metabolic rate per unit weight of rabbits increased consistently as the body weight decreased (p. 220).

The metabolic rate per unit surface area, however, was unaffected by body size, so Richet concluded that, for metabolic rate, surface area was more important than body weight. Simultaneously Rubner (1883 cit. by Krogh 1916, p. 133) noted a systematic decrease of the metabolic rate per unit weight of

fasting dogs as the weight increased from little dogs of 3 kg. to big specimens of 31 kg. body weight.

When the metabolic rate was expressed per square meter of body surface, however, the effect of body size disappeared. From this and similar observations, Rubner deduced his simple rule that fasting homoeotherms produce daily 1000 kcal. of heat per square meter of body surface.

Richet's and Rubner's intraspecific observation of the surface law was confirmed by interspecific comparisons. The best known table showing this surface law of metabolic rate is that published by Voit (1901). A 441 kg. horse produces over 948 kcal. daily per square meter of body surface, a 64 kg. man 1042, a 15 kg. dog 1039, and a 2 kg. hen 1008. So well established appeared the surface law that data which did not confirm it were either explained by particular conditions[1] or discarded as results of faulty measurements. Thus Lee (1929) suggests that Mitchell's rat surface is too high because it makes rats disobey the surface law.

THE "TRUE" BODY SURFACE AREA. Large and small bodies of similar shape have surface areas in proportion to the squares of their linear dimensions or the 2/3 power of their volumes. If the two bodies have also the same density, then their surface areas are also in proportion to the 2/3 power of their weights.

In this general way Sarrus and Rameaux understood the surface area of animals. Richet apparently preferred a definite surface that could be visualized. He calculated the surface area of his rabbits from their weights assuming they were spheres with a density of 1 kg. per liter. The surface area of such a sphere is $4.84\ W^{2/3}$ square decimeters when W is the weight in kilograms.

Considering that the animals are really not spheres, Meeh (1879), working under Vierordt, substituted in Richet's formula for the factor 4.84 a term, k, which is constant only within a group of similarly shaped animals but differs according to the shape of the animal.

By measuring animal surfaces and weights one may determine the Meeh constant empirically and then calculate the actual surface of similar animals as $S = K \cdot W^{2/3}$, where W is the body weight in kilograms and S the surface area in square decimeters. A set of Meeh constants is given by Lusk (1928, p. 123).

The surface of man averages 12.3 square decimeters per unit of the 2/3 power of body weight ($kg.^{2/3}$). For slim people the Meeh constant would be higher than this average, for stout people lower. Du Bois and Du Bois (1916) developed a formula which allows calculating the actual surface area of stout and slim human beings more accurately than does the Meeh formula.

Du Bois' formula reads:

$$S = 71.84\ W^{0.425} \cdot L^{0.725}$$

S = Surface area in cm.^{2}
W = body weight in kg.
L = body length in cm.

It is probably the best method of estimating the "actual" surface area of man. Du Bois' formula is dimensionally correct and is therefore valid for any size,

[1] Rabbits produced daily only 776 kcal. of heat per square meter; subtraction of their ear surface, however, brought their metabolic rate up to the more legal level of 917 kcal.

as it should be. (Dimensionally correct means that the surface area is expressed in terms of the 2/3 power of the volume or in the square of the linear dimension.)

Since for similar bodies L is proportional to the mean linear dimension, $W^{1/3}$, Du Bois' formula for comparison of large and small humans of the same relative stature may be written as

$S = k_1W^{0.425}\cdot k_2W^{1/3\times 0.725}$ which amounts to $k_1\cdot k_2\cdot W^{(0.425+0.242)}$ or $k_1k_2\cdot W^{2/3}$

Some later formulas for the calculation of surface area from the weight are dimensionally incorrect, such as the Moulton formula for cattle in which the surface area is proportional to the 5/8 power of body weight, and particularly the formula used by Ritzman and Benedict (1931) for calculating the surface area of sheep as proportional to the 0.561 power of body weight. The formula may have expressed the surface area of the sheep used for the measurements but it is theoretically incorrect for two animals of similar form which differ in size (but have the same surface per kg.$^{2/3}$). The Ritzman-Benedict formula would predict too small a surface for the larger of two animals with similar build.

Many methods have been invented for measuring the surface area of animals. In their eagerness to refine the surface measurements, many workers in this field seem to have overlooked a major question: "What is meant by the surface area?" Unless this question can be answered definitely, how can one decide which of two methods measures the surface more accurately?

One should obviously know whether or not the "true" surface of a rabbit is to include the surface area of the rabbit ears (see Benedict, 1934). As long as this question is open, which means an uncertainty of about 20 per cent, what is gained by refining the surface measurements to an accuracy of one per cent? According to Lee's (1929) (see p. 514) argument, the true body surface is the one that makes an animal's metabolic rate obey the surface law!

Kleiber (1932) pointed out that the various refinements of surface measurements and calculations, instead of clarifying issues, led to a more and more chaotic situation.

A great many published results of good work on metabolic rates are practically lost for comparative physiology because they are expressed only per unit of surface area, and the authors did not furnish the data which would make a comparison with other work possible. It was stated that the surface area was not well enough defined to serve as a basis for measurement, and, following Krogh's (1916) suggestion, and in agreement with Stoeltzner (1928) and Brody, Comfort and Matthews (1928), a power function of body weight was suggested as the basis of metabolic body size. The 3/4 power was proposed as the best fitting function (Kleiber, 1932).

Since then the situation has not improved. In the Annual Review of Physiology (Kleiber, 1944) alarm is again expressed as follows:

"In 10 papers (from 8 laboratories) studied for this review metabolic rates of rats are expressed per unit of the surface area. Four of the 10 authors did not state how they measured or calculated this area. One multiplied the 2/3 power of body weight (in kg.) by 7.42, another by 9.1, a third by 10, to calculate the

surface area in square decimeters. One author multiplied the 3/5 power of body weight by 12.44, and two have calculated a new surface-weight relationship, presumably by an intricate logarithmic interpolation between three older ones $7.42 \times W^{2/3}$, $11.36\ W^{2/3}$ and $12.44\ W^{3/5}$. That—for rat metabolism only—is this year's result of a century of surface law."

For man, the formula of Du Bois seems to be so generally accepted that the reports do not differ as widely as do those of rats, and clinicians may not have the time to develop new surface formulas for every new paper they publish. But also with human metabolism the valuable fruit of years of tremendous work cannot be used properly for comparative physiology, because the results are reported only in terms of the Du Bois surface without the data (height and weight) which would make a recalculation of the original results at least possible, even if extremely time consuming.

The comparisons of metabolic rates merely on the basis of surface area may be all a clinician wants. For comparative physiology more basic data are necessary.

In the interest of economy in research, physiologists should agree on minimum requirements for publication of metabolic data, and among those should be body weight and either total metabolic rate or metabolic rate per unit weight.

THEORETICAL VALIDITY OF THE SURFACE LAW. *A. The various theories of surface law.* The theories advanced for the interpretation of the surface law of animal metabolism may be classified into 5 major groups:

The metabolic rate of animals must be in proportion to their body surface.

1. Because the rate of *heat transfer* between animal and environment is proportional to the body surface area.

2. Because the intensity of flow of nutrients, in particular oxidizable material and oxygen, is a function of the sum of *internal surfaces* which in turn is proportional to the body surface.

3. Because the rate of supply of oxidizable material and oxygen to the tissues is a function of the mean intensity of the *blood current*, which is proportional to the square area of the blood vessels, which in turn is proportional to the area of body surface.

4. Because the *composition* of the animals is a function of their body size. The composition may be meant either *anatomically*: the larger the animal the lower is the ratio of the mass of metabolically active organs to the mass of metabolically inert organs; or the composition may be meant *chemically*: the larger the animal the lower its percentage of "active protoplasm".

5. Because the cells of the body have an *inherent requirement* of oxygen consumption per unit weight, which is smaller the larger the animal.

B. Critique of the theories of surface law. 1. The *heat transfer theory* is the most convincing of the five interpretations of the surface law. If homeothermism has survival value, and if the rate of heat transfer per unit surface area of large and small bodies is equal, then the surface law may be understood as the result of natural selection.

Accepting for the sake of the present discussion the biological advantages of maintaining a constant body temperature, one may ask: "How accurate is the

postulate that the rate of heat transfer per unit of surface area of large and small bodies is the same?"

The rate of heat flow of a body covered by an insulating layer may be formulated according to Fourier (Mach 1919, p. 84):

$$q = S\lambda \frac{T_i - T_s}{L}$$

where q = rate of heat flow; S = surface area
L = thickness of the covering layer (skin and fur)
T_i = internal temperature
T_s = temperature at the surface
λ = heat conductivity.

In animals the heat conductivity, λ, of the body covering, and its thermoconductive thickness, L, are rather hard to measure and are, furthermore, variable, since the depth of the layer with a temperature gradient is under vasomotor control. To simplify the problem, one may define the ratio $L/\lambda = r$ as the specific insulation of the animal covering, and write Fourier's law as follows:

$$q = S \frac{T_i - T_s}{r}$$

For a given difference between internal temperature and surface temperature, the rate of heat transfer is proportional to the surface area when the specific insulation for large and small bodies is the same. The specific insulation of animals is, however, variable.

The classical demonstration of this fact is the experiment of Hoesslin (1888). He reared two littermate dogs, one at 32°, the other at 5°C. The latter had to cope with a temperature difference between body and environment six times as great as the corresponding difference for his brother. Yet the metabolic rate of the dog in the cold was only 12 per cent higher. He solved the problem of keeping warm by growing a fur that weighed three times as much as that of his brother.

In the same animal and within a short time, the specific insulation may be reduced to 1/2 when the animal is transferred from a cold to a warm environment (Kleiber, 1932). Instead of maintaining the metabolic rate per unit surface area constant, large and small animals therefore might maintain a constant metabolic rate per unit weight, and with a variable specific insulation adapt the rate of heat loss to that metabolic rate.

The range within which the specific insulation can be changed is, however, limited. One may calculate (l.c., p. 327) that a 60 gram mouse with the same metabolic rate per unit weight as a steer, would need the equivalent of a steer's surface covering in a 20 cm. thick layer to maintain its body temperature in an environment of 3°C. This rather extreme example illustrates why it is advantageous for a small animal to have a higher metabolic rate per unit weight than a large animal. One can similarly show why at a high environmental temperature, for the prevention of overheating, it is advantageous for a large animal to have a lower metabolic rate per unit weight than a small animal.

Considerations of heat transfer thus make the assumption reasonable that natural selection would weed out such representatives of cattle and rats that insisted on having the same metabolic rate per unit body weight. These considerations however do in no way support the hypothesis that the metabolic rate of animals should be strictly proportional to their exact body surface, and that the results of metabolism studies would be more accurate and more reliable the more accurately one measured the "true" surface area of the animals.

2. The theory of *internal surfaces* as interpretation of the surface law would be valid only if the size of cells (or alveoli in the lungs as recently referred to again by Gaija, 1946) were proportional to animal size, if, in other words, elephants were made up of the same number of cells as mice. This logical prerequisite for the theory of internal surfaces does not seem to be supported by histology.

3. Attempts have been made to derive an interpretation of the surface law on the basis of *blood circulation*. The difference in the oxygen content of arterial and venous blood is independent of body size. The rate of oxygen consumption is therefore proportional to the intensity of the blood current, as already postulated by Sarrus and Rameaux (1839). This current intensity is the product of the linear velocity and the cross section area of the duct. Hoesslin (1888) declared that the cross section area of the aorta is proportional to the 2/3 power of body weight and the linear velocity is independent of body size, that consequently the intensity of the blood current is proportional to the 2/3 power of body weight, which is a measure for body surface. That the square area of the aorta should be proportional to the 2/3 power of body weight, appears sound application of the principle of similarity. Why, on the other hand, the same linear velocity of blood flow in the aorta for large and small animals should have particular survival value is not quite clear. (See Kleiber, 1932, p. 332.)

More convincing support of a circulatory interpretation of the surface law is gained by postulating, on the basis of the similarity principle, that heart volume or stroke volume of large and small animals should be approximately in proportion to their body weights.[2] On mechanical grounds the assumption seems reasonable that the heart of a horse cannot beat with the frequency of the heart of a mouse. The intensity of the blood current (stroke volume times frequency of heart beat), therefore, should increase less than in direct proportion to increase in weight.

$I = f \cdot v$, where I = intensity of blood current, f = frequency of heart beat and v = blood volume per heart beat. If the stroke volume is proportional to body weight, $v = kW$ then, $I = f \cdot kW$, and the Intensity per unit weight $\frac{I}{W} = kf$.

To explain the Surface law, the intensity of the blood current should be proportional to the 2/3 power of body weight, and, therefore, the frequency of heart beat should be inversely proportional to the cube root of body weight. $\left(f = k_1\frac{I}{W} = k_2\frac{W^{2/3}}{W} = k_2W^{-1/3}\right)$. From data given by Rihl (1927) for normal pulse frequency of various animals ranging from

[2] The similarity is actually not quite strict. Clark (1927, p. 72) concludes from interspecific comparison that heart weight varies with the 0.9 power of body weight. The heart weight per unit body weight has thus the tendency to become smaller the larger the animal.

rabbits to elephants, one may deduce the rule that the pulse frequency is inversely proportional to the 4th root of body weight. (The pulse frequency is on the average 186 $W^{-1/4}$ if W is given in kg.)

Considering heart volume and pulse frequency, one may therefore be satisfied that, from the point of view of oxygen transport, it is advantageous for large animals to have a lower metabolic rate per unit weight than small animals. Obviously, this consideration does not lend support to the hypothesis that the metabolic rate should be strictly proportional to the "true" body surface.

4. There appears to be very little evidence in support of the idea that the *chemical composition* of animals changes systematically with body size, so that the surface law can be understood on the basis of chemical composition. The concept of "metabolically active protoplasm" seems to be too vague for discussion. It seems that the activity of the protoplasm is determined by the metabolic rate; and what would then be gained by explaining, in turn, the metabolic rate on the basis of active protoplasm? The measurement of the nucleo-protoplasmic mass seems to offer an independent determination of active protoplasm, but Lindeman (1943) failed to find a correlation between nucleo-protoplasmic mass and rate of oxygen consumption of the retina.

The concentration of respiratory enzymes, especially cytochrome *c*, may be related to metabolic rate (Rosenthal and Drabkin, 1943), however this concentration itself, and particularly the activity of the enzymes, probably is under the influence of regulators, especially the endocrine system. (See Barron, 1943.) Enzyme concentration and enzyme activity may thus offer a valid explanation for the metabolic rate of tissues but in regard to body size and metabolic rate, enzyme concentrations and activity call for further explanation on the organismic level.

A good deal of speculation has centered around the anatomical composition as an explanation of the surface law.

Dreyer, Ray, and Walker (1910, p. 158) advanced the theory that the blood volume of animals was proportional to their surface areas. On the basis of the very data from which these authors draw this conclusion, one may demonstrate that, in general, the blood volume is more nearly proportional to body weight than to body surface (Kleiber, 1932, p. 329).

In direct opposition to the conclusion of Dreyer, Ray and Walker, furthermore, is a later observation by Gibson, Kelley and Pijoan (1938), who noted that in a series of dogs the blood volume per kg. body weight increased from 84 cc./kg. in 5 kg. dogs, to 97 cc./kg. in 30 kg. dogs.

Sweeping conclusions, on the basis of measurements on limited material, are often misleading.

Brody, Comfort and Mathews (1928, p. 33) made the following statement: "The weight of the kidney, the weight of the liver and practically the weight of the lung, blood, stomach, and intestine increase directly with the body weight at the same relative rate as does the surface."

This statement presumably led to the following sentence by Benedict (1938, p. 205):

"Our conclusion is, therefore, that if the brain weight is closely related to the heat production among animals other than the primates, this is in large part due simply to the general morphological development in proportion to the two thirds power of the body weight shown by innumerable body structures."

Before such generalizations are announced, they should be tested rationally by application to extreme cases. This can be done with the allegation that the blood volume is approximately proportional to body surface.

According to Lindhard (1926), the amount of blood in man is about 5 per cent of the body weight. If the statement of Brody et al. or Dreyer et al. were generally correct one would have to conclude from the blood content of a 70 kg. man that 49 per cent of the body of a 70 g. rat would be blood—indeed a bloody rat!

The following calculation similarly illustrates that one should be cautious when one derives a general quantitative rule from the tendency, that large animals have relatively smaller brains than the small animals.

A 5.4 ton elephant had a brain of 7.5 kg. (Benedict, 1936). The brain weight amounted, thus, to 0.14 per cent of the body weight. If the brain weight (in non primates) were proportional to the 2/3 power of body weight, the brain of a 5.4 gram shrew would amount to 14 per cent of its body weight, or that of a 100 gram rat to 5 per cent of its body weight, which is about 3 times as much as it actually weighs (Donaldson, 1924).

There seems to be a general tendency for the larger animals to economize on brain weights, as J. B. S. Haldane (1946) pointed out in his stimulating essay ("On being the right size"), but what he observes within a group of Felidae from cat to tiger can not be regarded as a general quantitative rule. If generally the weight of the brain were only doubled whenever the body weight is quadrupled, or in other words, if the brain weight were proportional to the square root of body weight, then, starting down from the elephant with a brain weight of 0.14 per cent of its body, one would conclude that one-third of a 100 gram rat should be brain. Similarly, if Dubois' (cit. by Weber, 1923, p. 151) formula, making brain weight a function of the 5/9 power of body weight, were generalized, the 100 gram rat would have to have a brain weight of 18 grams. Neither Dubois nor Haldane can be accused of generalizing the brain weight-body weight relationship so much as to include the comparison of rats and elephants. Such generalization is, however, made when this relation of brain weight to body weight is used to explain the surface law of metabolic rate, because that law includes comparisons of rats with elephants.

The idea that the surface law of metabolism can be explained by the anatomical composition of the animal has been advanced especially by Blank (1934), and by Kestner (1934 and 1936) who made the following statement. "Hence I conclude that the relative sizes of the brains and the large glands can give a complete explanation of the different heights of metabolism of different animals."

In support of this idea Kestner gives the ratios of brain weight to body weight of large and small dogs. A dog of 40 kg, for example, weighs 400 times as much as its brain. A dog of 5 kg, however, weighs only 25 times as much as its brain.

By calculating from these data the absolute weight of the brains we would be led to the somewhat surprising result that the brain of the small dog weighs 200 grams; that of the large dog only 100 grams.

To support his point that higher metabolic rates of the smaller animals are explainable by their greater relative brain size, Kestner states that the brain weight of a 3 kg. dog is about 50–60 grams whereas that of a rabbit of the same body weight is only 10 grams. This comparison, however, far from supporting Kestner's theory that the brain weight dominates the metabolic rate, to the contrary is very good evidence against this theory. Despite the great differences in relative brain weight noted by Kestner, the mean metabolic rate of rabbits is only about 10 per cent lower than that of dogs of the same size (see table 1).

It is generally accepted that primates have exceptionally large brains. If brain weight dominated metabolic level, primates should have a particularly high metabolic rate. Yet a group of college women with an average body weight of 54.8 kg. (and an average brain weight of probably not much less than 1300 grams) had an average basal metabolic rate of 1224 kcal. per day, which is slightly less than the corresponding average of 1254 kcal. per day for a group of female sheep with an average weight of 46.4 kg. and a brain weight of, presumably, about 100 grams (Pálsson 1940).

TABLE 1

	BODY WEIGHT	DAILY HEAT PRODUCTION PER SQUARE METER	
		1901	1931
	kg.	*kcal*	*kcal*
Horse	441	>948	
Steer	342		1465
Man	64	1042	926
Dog	15	1039	776
Hen	2	1008	676

5. ***Body size and tissue metabolism*** is a problem that is not yet satisfactorily solved. In the paper in which Rubner (1883) demonstrated the surface law in the metabolic rate of dogs he wrote (p. 550):

"Large and small dogs have a different metabolic rate, not because there are definite differences in the organization of their cells, but because the impulses originating in the skin from cooling stimulate the cells to metabolic activity."

The sum of these impulses, according to Rubner, is in proportion to the body surface.

Later on Rubner accepted just the opposite view, declaring that the metabolic rate of the tissues is a fundamental characteristic of a species (Wels, 1925).

In 1925 Terroine and Roche derived from microrespiration trials with excised animal tissue a hypothesis opposed to Rubner's idea of genetically fixed cell metabolism. They stated: "Homologous tissues of different animals have, in vitro, the same intensity of respiration." The metabolic differences of tissues

of large and small animals disappear when the tissues are removed from the animal body. These differences, therefore, are not located in the cells themselves, but are imposed on the cells by central systems representing the organism as a whole, the nervous system and the circulatory system.

In the same year Grafe (1925), independent of Terroine and Roche, reached the same conclusion stating: "The living protoplasm of the warm blooded animals and maybe even of many cold blooded animals, shows, as far as respiration is concerned, a certain uniformity, and gets its specificity only by means of the regulating system of the animal!"

The tissues, according to Grafe, respire at a higher rate in vitro than in vivo, particularly tissues of large animals. In vivo the metabolic rate of those tissues is checked by the influence of the central regulators, mainly the nervous and endocrine systems.

There are a number of observations in support of this view. Severance of the spinal cord increased the rate of nitrogen metabolism (Isenschmid, 1920). In line with these ideas is also the observation of Victor (1934) that dystrophic muscles have increased metabolic rate in vitro. Recently Houchin (1942) stated that normally tocopheryl phosphate acts as a brake on muscle oxidation, and that vitamin E deficiency lets the oxidation "run riot" so that the muscle is "consumed in its own fire" and muscle dystrophy results.

The discovery of capillary regulation of blood supply to the tissues by *Krogh* (1929) led to a well rounded theory that the checking of metabolic rate in vivo was accomplished by rationing of the oxygen supply. That was in contrast to Pflüger's law (Pflüger, 1872), to be sure, but observations of Thunberg (1905) on lower animals, and of *Verzár* (1912) on muscles, indicated that Pflüger's law loses its application at low partial oxygen tensions and that oxygen supply may become the limiting factor in the rate of tissue respiration. Recently Gaija (1946) noted that below a given critical partial pressure of oxygen the metabolic rate of homeotherms depends on oxygen supply. Apparently, independent of body size, the limiting partial oxygen pressure amounts to about 90 mm. Hg. This corresponds to a barometric pressure of 450 mm. Hg. or an altitude of 4000 meters.

That the metabolic rate of animals was by no means a genetically fixed constant, but was decidedly affected by somatogenic factors, was demonstrated by Kleiber and Cole (1939) who measured the metabolic rate of giant rats produced by injection of growth hormone. The metabolic rate, either per unit weight or per unit of the 2/3 power of weight, was lower in the growth hormone rats than in normal litter mate controls. This rate could therefore not be a genetically fixed characteristic of the tissues (being changed by the somatic influence of growth hormone injection). This result was in line with the ideas of Terroine and of Grafe. Kleiber and Cole, however, observed also that the differences in the metabolic rates of giant growth hormone and normal control rats could still be noted in the rate of oxygen consumption of the surviving diaphragm in vitro. This observation was in opposition to the results of Terroine and of Grafe.

Field, Belding and Martin (1939) succeeded in summating the metabolic

rates of excised rat tissues in vitro to 66 per cent of the actually measured metabolic rate of the rat. Martin and Fuhrman (1941) similarly summated metabolic rates in vitro of dog tissues to over 70 per cent of the metabolic rate of the dogs. These results support the idea that the rate of oxygen consumption of the tissues in vitro reflects the metabolic rate of these tissues in vivo to a considerable extent.

Repeating the investigations of Terroine and of Grafe, Kleiber (1941) measured the in vitro rate of oxygen consumption of liver slices from rats, rabbits, and sheep. The metabolic rate per unit weight in vitro decreased consistently with increasing size of the animals, almost to the same extent as the metabolic rate of the living animal.

Recent microrespiration trials in our laboratory (Kleiber, 1947, unpublished) with excised liver of rabbits and cows again show that liver slices from cows have a lower rate of oxygen consumption per unit weight than the liver slices from rabbits.

Smyth (1940), on the other hand, reports that "observations by Van Heyningen (1936) indicate a uniformity in the rate of metabolism in the same tissue in different species irrespective of the size of the animal."

This result seems to be in line with the theory of Grafe and of Terroine and in contrast to our findings. A check of Van Heyningen's figures, however, indicates that only anaerobic glycolysis of brain cortex was independent of body size. For respiration, the effect of body size on tissue metabolism was similar to our results.

The following working hypothesis seems justified: The animal as a whole responds to changes in somatic conditions. Reactions to such changes are transmitted to the tissues by centralized metabolic regulators such as the nervous and the endocrine systems. These regulatory influences produce changes in the metabolic conditions of the cells. Among these conditions, which determine cellular metabolic rate, may be structural relations, electrical potentials, ion concentrations, and the concentrations of oxygen, metabolites and enzymes.

The observation that metabolic rates in vitro are similar to those in vivo suggests that the conditions which determine the rate of cell respiration change but slowly; this would be expected of changes in concentrations of enzymes and metabolites.

Since oxygen supply should not affect the oxidation rate in vitro, the similarity between in vitro and in vivo metabolic rate throws doubt on the theory that tissue metabolism normally is regulated by rationing the oxygen supply.

The observation that the metabolic rate of tissues in vitro is similar to the corresponding rate in vivo, is sometimes interpreted to indicate an inherent metabolic level of the tissues. If the statement means that the rate of tissue respiration is independent of somatic influences, then it is wrong; if it means that genetic factors are involved in the metabolic levels, then the correctness of the statement is hardly in doubt. One may be able to breed strains of rats with a high, and strains of rats with a low metabolic level under quasi equal environmental conditions. There may be genetically determined limitations

for the extent to which cell metabolism reacts on changes produced in the cells by the metabolic regulators. The demonstration of genetic factors in metabolic level, however, helps but little in the interpretation of the fact that the metabolic rate per unit weight of a 5.3 kg. rabbit is practically the same as that of a 6.6 kg. dog, but differs significantly from that of a 1.5 kg. rabbit; as in turn the metabolic rate per unit weight of a 25 kg. dog differs significantly from that of the 6.6 kg. dog.

The results of Terroine and of Grafe on metabolic rates in vitro were not confirmed later, yet an essential part of their theory is sound; namely, the idea that the metabolic rate of the tissues in vivo is controlled by central regulators representing the animal as a whole.

Even fetal metabolic rate seems to be under the influence of the maternal metabolic regulators (or perhaps a combination of maternal and fetal regulators). Kleiber, Cole and Smith (1943) measured the metabolic rate of rat embryos in vitro, and obtained good evidence against the theory that a fetus behaves, metabolically, like an independent small homeotherm. The fetal metabolic rate per unit moist weight was of the same order of magnitude as that of normal adult rats, and considerably smaller than that of newly born or 12 day old rats.

This observation, like others mentioned, favors the idea that, to a considerable degree, cellular metabolic rate is adapted to the condition of the animal as a whole. The effect of the metabolic regulators, however, appears to remain in the tissues that are removed from the animal and respire in vitro.

The analysis of these factors controlling the metabolic level in vitro and in vivo appears to be a most fruitful field for future research on tissue metabolism.

C. Integration of valid theories on surface law. Of the five groups of interpretations of the surface law, the one on internal nutritive surfaces has no evidence in its favor. The one on composition of the body has no explanatory value. The same is true for the explanation with a genetically fixed metabolic rate of body cells, even if this postulate as such were acceptable, because the problem would arise: why have those animals been selected for survival whose cells have the inherent metabolic rate such that the metabolic rate follows the surface law?

The theories that relate the surface law to rate of heat transfer, and those that relate it to the hemodynamics, have most value for the interpretation of the surface law. The most promising theory of body size and metabolism integrates the heat exchange and the circulation theory.

In natural selection, those animals probably prove to be the fittest whose cells are adapted to such a level of oxygen consumption that the metabolic rate of the animal is most suitable for the maintenance of a constant body temperature and in line with the most efficient transport of oxygen.

A four ton animal, whose cells insisted on a rate of oxygen consumption per unit weight equal to that of mouse cells, could not survive, because such a metabolic rate could not be supported by the circulatory system and would upset the maintenance of a constant body temperature.

The surface law is unreliable mainly because the definition of an animal's surface is vague. But even if the surface area could be defined and measured

accurately there is no theoretical basis for the hypothesis that the metabolic rate of homeotherms should be exactly proportional to their particular surface area rather than to a more general function of body size.

EMPIRICAL VALIDITY OF SURFACE LAW. Between the efforts of the gadgeteers to design apparatuses for surface measurements, the statisticians to derive formulas for calculating "true" surface areas, and the theoretically inclined biologists who discussed the proper interpretations of the surface law, rather few seemed to have been interested in the question as to just how reliable the surface law itself is.

Among the few that questioned rather early the validity of the surface law was one of the pioneers of vitamin research, F. G. Hopkins, who in 1912 wrote as follows:

"... in the case of very young rats the demand for maintenance is more nearly determined by the live-weight than by the surface area. It becomes of course smaller relatively to body weight as this increases but it falls off more slowly than the surface area would require."

Seven years later, Harris and Benedict (1919) concluded from their extensive measurements on men and women that within the human species there was no evidence for the surface law. Du Bois (1927, p. 202) however, maintained that the data of Harris and Benedict confirmed the surface law.

The differences in size among the material of Harris and Benedict were so small and the influence of factors other than size so relatively large that the two opposing deductions could both be made in good faith (compare pp. 530 and 534).

In an attempt to find the most suitable unit of metabolic body size for estimating maintenance requirements and comparing metabolic rates of animals that differ in weight, Kleiber (1932) compiled results of metabolism measurements from American laboratories, listing 13 groups from 150 gram ring doves to a 679 kg. steer. The surface law was confirmed insofar as the metabolic rate per unit weight decreased systematically as the body weight of the animals increased. The coefficient of variability per unit weight was 80 per cent, and that per unit surface area only 34 per cent. This latter is still quite high, and a comparison of some of the recent results with Voit's table (1901) apparently indicates a trend of the modern American animals to take the surface law less seriously than the European animals in 1901.

There is a considerable positive correlation between body size and metabolic rate per square meter of body surface.

The surface law, that is the theory that the metabolic rate per unit surface area of large and small animals is the same—or at least independent of size—is therefore not strictly confirmed by the recent material.

METABOLIC RATE AS A POWER FUNCTION OF BODY WEIGHT. *A. Linear relation between logarithms of metabolic rate and body weight.* The regularity in the deviations of the empirical results from the surface law justified the search for a function of body size to which metabolic rate might be more nearly proportional than to body surface. Plotting the logarithms of fasting metabolic rate against the logarithms of body weight revealed a linear relation between these two variables

with surprisingly small deviations from the mean trend. For the ten groups of mammals, the standard deviation from the mean regression line amounts to ± 0.03 logarithm unit, which corresponds to a coefficient of variation of metabolic rate from the interspecific mean of 7 per cent.

When the logarithm of metabolic rate is a linear function of the logarithm of body weight, then metabolic rate is proportional to a given power of body weight.[3] The metabolic rate was more nearly proportional to the 3/4 power of body weight than to either the 2/3rd power or the surface area of the animals (as derived for each group by a special formula designed for the calculation of its particular surface area).

Soon after the publication of these results, Brody and Procter (1932) extended a similar compilation of metabolic data down to the mouse. They used mostly figures accumulated in their own laboratory and noted that their result was in close agreement with that of Kleiber (l.c., p. 94).

In 1938, F. G. Benedict published an extensive analysis of the great amount of valuable data on metabolic rates of various animals tested in the Carnegie Nutrition Laboratory by reliable methods under rather well standardized conditions.

A chart (p. 171) shows the logarithms of the mean metabolic rates for 28 groups of animals plotted against the logarithms of the corresponding mean body weights. The animals considered ranged in size from 20 gram mice to nearly 4 ton elephants. A regression line indicates the average trend, and Benedict notes "a most gratifying straight line relationship between the total heat production and the body weight." He obviously means the logarithms of these variables. He then continues as follows: "However satisfactory this relationship may be mathematically, this method of presenting the data completely masks metabolic differences within species."

In answer to this remark one may say that any mean of a group of data "masks," that is, does not show, the differences between the single data; that has nothing to do with logarithmic interpolation. If one wants to show the differences within the species one may plot the logarithm of each individual metabolic rate against logarithm of the corresponding individual body weight. In order to supply the reader with an estimate of the variability, as a substitute for presenting every single result, some of the more statistically minded biologists state with a mean also its standard error.

Benedict extends his accusation, stating that logarithmic interpolation "distorts or obscures striking differences between the species." Since, however, a logarithmic chart in a scientific paper is presented to readers who are presumably familiar with logarithms, the accusation of distorting or obscuring can be discarded.

There is nothing obscure about the fact that a logarithmic regression line of a given set of data looks different, in general, from the corresponding arithmetic line, and if this difference in the appearance of the two regression lines be termed

[3] If $\log M = \log a + p \cdot \log W$—(general equation for straight line)
then $M = a \cdot W^p$

distortion, one could call the arithmetic line a distortion of the logarithmic just as well as vice versa.

Benedict further believes (p. 172) that "the seeming similarity between the different species shown by this logarithmic chart is an artificial similarity." It is true that the conditions under which basal metabolism is measured are somewhat different from those under which animals normally live. The conditions under which the results discussed here are obtained may therefore be classified as artificial, and it might be argued that the metabolic rate of a rat in a sewer, or a cow in a barn, would be different from their metabolic rates measured in the artificial environment of a respiration apparatus; but that possibility has nothing to do with the question, whether or not the linearity of the logarithms of metabolic rates as measured and the corresponding body weights are real. Whether or not such a relationship has physiological significance depends on the physiologist, the same as it depends on the listener whether he hears a symphony or merely a multitude of different sounds.

Benedict concludes his chapter on interspecific comparisons (p. 179) with the following sentence:

"It seems, therefore, unjustifiable to apply mathematics to the pooled end result of the activities of millions of cells each highly differentiated, with different energy potentialities and actuated by different stimuli."

If this is the way Benedict feels, one cannot help but wonder how he ever became interested in conducting a respiration trial and why, furthermore, he even calculated means of groups of several of these pooled end results which indeed is applying mathematics.

Admittedly one may find biological publications in which data seem to be used merely as material for mathematical exercises. Admittedly, furthermore, some biologists learn mathematical tricks and tirelessly apply them, apparently without bothering to understand either the basis and limitations of their methods or the meaning of their results. Publications originating under such circumstances may have developed Benedict's antagonism to the application of mathematics in biology. The answer to such pseudomathematics in biology, however, is not less mathematics, but good mathematics.

For the application of statistical methods the biologist finds an excellent review by H. L. Dunn (1929) with ten "don't's" and two "be sure of's." These twelve commandments ought to be instilled into every student of quantitative biology and physiology.

Benedict's table 4 on p. 175 (1938) contains an error which may lead to some confusion and is therefore mentioned here. The metabolic rates of the animals lighter than 1 kg. seem to have been calculated per unit of the various power functions of body weight: $W^{0.62}$, $W^{2/3}$, $W^{0.72}$ etc., on the basis of the erroneous assumption that generally (weight in g.)^p equals 100 (weight in kg.)^p. For $p = 2/3$ the result happens to come out all right because $1000^{2/3} = 100$ but for all the other powers that calculation is wrong; obviously it would be wrong for $p = 1$, since 1 kg. contains 1000, not 100 grams. The one kcal. of heat produced daily by the 8 gram dwarf mouse amounts to 39.2 kcal. per kg.$^{0.76}$ not 20.6 kcal. as listed in the table mentioned.

The great amount of material on metabolic rates, secured by reliable measurements in Benedict's laboratory, and condensed to a logarithmic chart on the relation of body size and metabolic rate (Benedict (1938) p. 171) is an excellent confirmation of Kleiber's earlier results (1932, p. 321). The two regression lines expressing the mean trends of these two sets of data are practically identical.

B. Check with recent data. To reinvestigate the relation of body size and metabolic rate among mammals, I have compiled 26 groups of metabolic rates measured under apparently comparable conditions. The animals were mature, in postabsorptive condition, measured in the range of metabolically indifferent environmental temperature, and at rest, or at least without abnormal activity. No data were used that were already incorporated in the earlier study (1932).

The basic data and the source of the material are given in table 2.

In figure 1 the logarithms of metabolic rate are plotted against the logarithms of body weight. The results used for calculating the regression line are indicated by dots in circles; those that are not used in the calculation are marked with brackets.

The reasons for not including these data in the calculation are as follows:

The result of only one shrew is so far reported and it is questionable whether the conditions of measurement allow a direct comparison with standard metabolic rate of the other animals. The Swiss mice were not in postabsorptive condition as indicated by a mean respiratory quotient of 0.96. Dwarf mice and growth hormone rats have an abnormal endocrine system. The result for swine was calculated from a mean net energy requirement for maintenance. The steer calves may be regarded as not yet mature and therefore their metabolic rate not strictly comparable with that of the other groups of animals. The conditions of measurement of metabolic rate of elephant, porpoise and whale are not strictly in line with the normal conditions of measuring standard metabolic rates.

Unfortunately, a lot of valuable data on man could not be incorporated in our chart because the results were given only per square meter of body surface. This is particularly true for the material of Boothby and Sandiford (1924), and the more recent findings of Young, Pittman, Donelson, and Kinsman (1943).

For the 26 comparable results, the method of least squares leads to the following linear regression equation:

$$\log M = 1.83 + 0.756 \log W \pm 0.05$$

where M = metabolic rate of animal in kilocalories per day

W = body weight in kilograms.

The regression coefficient of 0.756 ± 0.004 indicates that for the 26 results compiled in our table, the metabolic rate is most nearly proportional to the 0.756 or close to the 3/4 power of body weight.

The same interpolation of the earlier data (Kleiber, 1932) on ten groups of mammals leads to the equation:

$$\log M = 1.87 + 0.739 \log W \pm 0.03$$

TABLE 2

GROUP	ANIMAL	AUTHOR	BODY WT.	METABOL. RATE PER DAY
a. Data used for calculation of regression line				
			kg.	kcal
1	Mouse	Benedict and Lee, 1936	0.021	3.6
2	Rat 230–300 days old	Kleiber, unpubl.	0.282	28.1
3	Guinea pig	Benedict, 1938	0.410	35.1
4	Rabbit	Tomme and Loria, 1936	2.98	167
5	Rabbit	R. Lee, 1939	1.52	83
6	Rabbit	R. Lee, 1939	2.46	119
7	Rabbit	R. Lee, 1939	3.57	164
8	Rabbit	R. Lee, 1939	4.33	191
9	Rabbit	R. Lee, 1939	5.33	233
10	Cat	Benedict, 1938	3.00	152
11	Macaque	Benedict, 1938	4.2	207
12	Dog	Galvão, 1942	6.6	288
13	Dog	Galvão, 1942	14.1	534
14	Dog	Galvão, 1942	24.8	875
15	Dog	de Beer and Hjort, 1938	23.6	872
16	Goat	Benedict, 1938	36.0	800
17	Chimpanzee	Bruhn and Benedict, 1936	38.0	1090
18	Sheep ♀	Lines and Peirce, 1931	46.4	1254
19	Sheep ♂	Lines and Peirce, 1931	46.8	1330
20	Woman	McKittrick, 1936	57.2	1368
21	Woman	Lewis, Iliff and Duval, 1943	54.8	1224
22	Woman	McCrery, Wolf and Bavousett, 1940	57.9	1320
23	Cow	Benedict and Ritzman, 1935	300	4221
24	Cow	Kleiber, Regan and Mead, 1945	435	8166
25	Beef heifers	Kleiber, Goss and Guilbert, 1936	482	7754
26	Cow	Benedict and Ritzman, 1935	600	7877
b. Data not used for calculation because conditions not comparable				
	Shrew	Morrison and Pearson, 1946	0.0035	2.9
	Swiss mice	U. S. Navy Res. Unit and Kleiber, 1944	0.0105	3.7
	Dwarf mouse	Benedict, 1938	0.008	1
	Rats (giant)	Benedict, 1938	0.400	33.2
	Rats (growth hormone)	Kleiber and Cole, 1939	0.391	28.6
	Swine	Breirem, 1936	150	2678
	Steer calves	Mitchell et al., 1940	200	3817
	Elephant	Benedict, 1938	3672	49000
	Porpoise	Irving et al., 1941	170	6768
	Whale	Irving, 1941	70000	1.2×10^6

The two bands, described by the two regression equations with their standard error of estimate, could be distinguished only by extrapolation to animals weighing less than 4 grams or more than 800 tons.

The figure shows the line by which the results would have to be represented if the metabolic rate were proportional to body weight, and also the line which would summarize the results if the metabolic rate were proportional to the 2/3 power of body weight (or approximately body surface).

The line expressing proportionality of metabolic rate to body weight stays within the band expressing metabolic rates proportional to the 3/4 power of

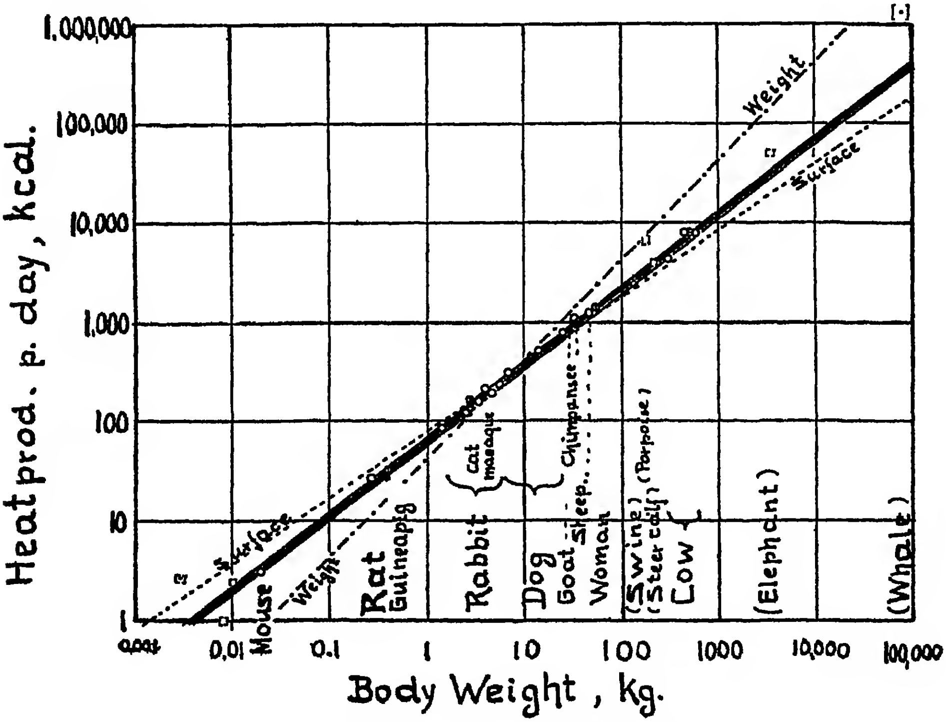


Fig. 1. Log. metabol. rate/log body weight

weight over a ratio of weights of 3.2. That means one would not be able to decide whether metabolic rates are proportional to body weight or proportional to its 3/4 power, unless the heaviest animals studied would weigh at least three times as much as the lightest animals.

Similarly, one may calculate that a significant difference between proportionality to the 3/4 power of body weight and proportionality to the 2/3 power of body weight (representing roughly the surface area), could not be established with groups of animals in which the heaviest animals weighed less than 9 times as much as the lightest animals.

C. Intraspecific comparisons. Since the differences in size have to be so con-

siderable to allow a significant decision of the question, whether the metabolic rate is more nearly proportional to the 2/3 or the 3/4 power of body weight, only a few species show great enough differences in mature weight for the establishment of the best fitting power function of body weight as basis for metabolic body size. This is true, even though the variability of metabolic rate for a given size is less within a species than between species.

In three species, namely, mice, rabbits, and dogs, the differences in size are considerable, and an analysis of the relation between body size and metabolic rate within these 3 species seems more promising than in the other species with more uniform size.

The regression line given for metabolic rate of mice (Benedict, 1938, p. 65) as a function of their body weight, would indicate a best fitting power function of body weight for metabolic body size

$$p = \frac{\Delta \log \text{metabol. rate}}{\Delta \log \text{body weight}} = 0.89$$

Benedict's line seems, however, definitely affected by a group of 4 endocrinologically abnormal dwarf mice. If we estimate the line through the other results, the best fitting power function would be 0.76.

The regression line given for rats (l.c. p. 67) would indicate that the rats intraspecifically followed the surface law. Their metabolic rate would be most nearly proportional to the 0.67 power of body weight.

The results of the 5 groups of rabbits reported by Lee (1939) may be represented by the following regression line:

$$\log M = 1.763 + 0.82 \log W \pm 0.09.$$

This indicates that in this population, metabolic rate is most nearly proportional to the 0.82 power of body weight; the range of body size in these rabbits would, however, have to be about 4 times as great as it is, to demonstrate a significant departure of rabbits from the hypothesis that metabolic rate is proportional to the 2/3 power of body weight. This range would have to be considerably more than four times as large to demonstrate that the intraspecific relation between metabolic rate and body size in these rabbits contradicts the hypothesis that the metabolic rate is proportional to the 3/4 power of body weight.

The data on the 3 groups of dogs reported by Galvão (1942) indicate proportionality of metabolic rate to the 0.84 power of body weight, and since the variability in this case seems rather small, the deviation from the surface law and even from the 3/4 power rule appears significant; so does the deviation from the hypothesis that the metabolic rate of these dogs is proportional to their body weight.

The result of Galvão contradicts the one which can be calculated on the basis of the material of Lusk, Kunde, and Steinhaus (see Benedict 1938 p. 79). Selecting from the regression line of this material, given by Benedict, the points for 9 kg. dogs and that for 18 kg. dogs, one may conclude that the metabolic rate of these dogs is most nearly proportional to the 0.55 power of body weight.

It is of interest to note that within the two groups of animals in which an intraspecific investigation of body size and metabolic rate appears most promising, rabbits and dogs, the metabolic rate seems to be most closely related to a power function of body weight higher than the 3/4. This problem deserves more study. Until there is a great deal more evidence for a difference in the intraspecific and the interspecific relation of body size and metabolic rate, the most rational temporary hypothesis is the one which is the simplest. That is the hypothesis that the intra- and the interspecific relations of body size and metabolic rate are the same. The question is whether the relation found intraspecifically should be assumed to hold also interspecifically, or vice versa. The interspecific derivation of the best suitable unit of metabolic body size seems, at present, the more reliable, because the great difference in size available by interspecific comparison makes size so dominant over other factors.

At present it appears, therefore, reasonable to adopt for intraspecific prediction of the metabolic rate the metabolic unit of body size found most suitable for interspecific prediction, namely, the 3/4 power of body weight.

Metabolic body size and prediction of metabolic rate. The metabolic body size is that function of size of animals to which the metabolic rate is proportional. Once the unit of metabolic body size is established, then the metabolic rate can be predicted by multiplying the metabolic body size with a given factor.

According to the surface law, the metabolic rate is proportional to the surface area. The metabolic body size, consequently, would be expressed in square meters of body surface, and according to Rubner's rule, the daily metabolic rate in kilocalories is 1000 times the number of square meters of body surface.

Since body surface is ill defined, the square meter of body surface is no suitable unit of size, even though it undoubtedly offers the great advantage of being easily visualized. Admittedly, it is easier to talk of heat production per square meter than to talk of heat production per unit of the 3/4 power of body weight in kilograms.

The Du Bois surface (1916) as calculated from weight and height is well defined within one species, but for comparison between man and rabbit, for example, one would have to rely again on the ill defined concept of "true" body surface.

Attempting to avoid hypotheses, such as the surface law, and ill defined terms, such as the body surface area, Harris and Benedict (1919) derived their empirical prediction equations, namely, (p. 227):

for men $h = +66.4730 + 13.7516w + 5.0033s - 6.7550a$

for women $h = +655.0955 + 9.5634w + 1.8496s - 4.6756a$

h = total heat production in kcal. per day, w = weight in kilograms

s = stature in centimeters and a = age in years.

Krogh (see Boothby and Sandiford, 1924, p. 80) has already criticised these equations.

That they allow an accurate prediction of the metabolic rate of men and

women similar to the material studied by Harris and Benedict is not in doubt, but physiologically the equations are practically meaningless.

Implied in these equations are several rather questionable assumptions; namely, that the metabolic rate of a weightless woman is 10 times that of a weightless man.

Further, it is assumed that metabolic rate is the sum of a linear function of body weight, a linear function of height, and a linear function of age.

Even though the Harris-Benedict equation deals with well defined quantities, such as weight, height, and age, Du Bois' prediction, despite its connection with the ill defined concept of body surface, seems to be more acceptable to clinicians as well as physiologists because it has a rational physiological meaning. Assuming that the metabolic rate is proportional to the 3/4 power of body weight, Kleiber (1932) re-analyzed the data of Harris and Benedict and developed the following prediction equations:

for men: $M = 71.2 \times W^{3/4} (1 + 0.004 (30\text{-}A) + 0.010 (S\text{-}43.4))$

for women: $M = 65.8 \times W^{3/4} (1 + 0.004 (30\text{-}A) + 0.018 (S\text{-}42.1))$

M = the metabolic rate in kilocalories per day

W = the body weight in kilograms

A = the age in years

S = the specific stature in $\dfrac{\text{centimeters height}}{W^{1/3}}$.

The equations have the following meaning: the metabolic rate of a man 30 years of age and with a specific stature of 43.4 cm./kg.$^{1/3}$ has a metabolic rate 71.2 times his body weight in kilograms raised to the 3/4 power. A woman of a standard age of 30 and standard specific stature for women (which is 42.1 or slightly less than that for men) has a metabolic rate of 65.8 per kg.$^{3/4}$. The ratio of 71.2 to 65.8 for the metabolic rates per unit of the 3/4 power of body weight of standard man and standard woman indicates the effect of sex on human metabolic rate. The second term in the parenthesis indicates that in adult man the metabolic rate decreases about 0.4 per cent of the metabolic rate of standard man for each year above (or increases for each year below) the standard age of 30.

It might have been preferable to express this age effect in relative terms, such as $0.12 \dfrac{30\text{-}A}{30}$. Then the metabolic effect of age, expressed by the coefficient 0.12, would be more comparable with that in other animals with different standard age. (The relation of relative ages in different animals has been extensively studied by Brody (1945, chapter 19). Human metabolic rate would change 12 per cent per unit of relative departure from standard age (that is per 30 years). The last term in the parenthesis indicates that the mean specific stature (in the material of Harris and Benedict) for men is 43.4 cm. per kg.$^{1/3}$ that for women 42.1 cm. per kg.$^{1/3}$. Each centimeter per kg.$^{1/3}$ increase in specific stature produces, on the average, an increase of 1 per cent of the metabolic rate of men, and of 1.8 per cent of the metabolic rate of women. This suggests that the degree of slender-

ness (specific stature) affects the metabolic rate of women more than it affects the metabolic rate of men. The equations of Kleiber predict the metabolic rate of human beings with about the same accuracy as the empirical multiple regression equations of Harris and Benedict. The major advantage of the newer equations is that all terms have a physiological meaning.

Kleiber (1932) was somewhat vague about choosing a definite power function of body weight as the best basis for metabolic body size, and stated, p. 336, "that further investigation may show that some unit other than $W^{3/4}$ may be preferable."

Brody and Procter (1932) ventured a more definite formulation, namely:

$$M = 70.4\ W^{0.734}$$

and Brody, Procter and Ashworth (1934) obtaining exactly the same equation with a wider range of animals, including elephants, were confident that this "equation approaches closely the true relation between basal metabolism and body weight of mature mammals."

This relation was widely accepted (for example, in Höber's book, *The physical chemistry of cells and tissues*, p. 375). The National Research Council conference on energy metabolism (1935) endorsed the power function, after reducing it to two decimals (0.73), as the most suitable unit of metabolic body size. There was some doubt as to whether or not the data available justified the change from the 3/4 to the more complicated 0.73 power, a change which implied that the second decimal of the exponent was significant.

Taking into account a coefficient of variation in metabolic rate per unit size of 7 per cent, one may calculate that there is no significant difference between Brody's prediction and $72W^{3/4}$ within a group of animals ranging from a 10 gram mouse to a 16 ton super elephant. Within this range, there is thus not much point in discussing the question, whether the 0.734 power or the 3/4 power of body weight fits the metabolic results more closely.[4]

Recently, Brody (1945, p. 373) dropped also the second decimal from the exponent and suggests that the 0.7 power be used as reference for basal metabolism measurements.

Our recent analysis favors again the 3/4 power of body weight as the metabolic unit of body size. Aside from its empirical justification, the 3/4 power is preferable to either of Brody's successive 0.734, 0.73 or 0.7 because it is mathematically simpler since it can be calculated without logarithms.[5]

The daily fasting heat production per kg.$^{3/4}$ was:

in 10 groups of mammals, 1932: 71 $\pm$ 1.8 kcal. per kg.$^{3/4}$
in 26 groups of mammals, 1947: 69 $\pm$ 1.5 kcal. per kg.$^{3/4}$
together:
36 groups of mammals, 69 $\pm$ 1.2 kcal. per kg.$^{3/4}$

[4] The classification of the 3/4 power of body weight as the "Brody-Kleiber unit" (Günther, 1944) nevertheless is confusing. If these units have to be named according to authors at all, then Brody's unit would be the 0.734, 0.73 or 0.7 power, and Kleiber's the 3/4 power.

[5] $W^{3/4}$ can be easily obtained on a slide rule (extract the square root of the square root of the cube of the body weight).

For all practical purposes, one may assume that the mean standard metabolic rate of mammals amounts to 70 times the 3/4 power of their body weight (in kg.) per day, or about 3 times the 3/4 power of their body weight (in kg.) per hour.

BODY SIZE, FOOD REQUIREMENT AND DOSAGE OF BIOTICS. Two animals may be regarded as being on the same level of food intake when their rate of intake of metabolizable food energy[6] is the same multiple, or the same fraction, of their standard metabolic rate. Since the standard metabolic rate is proportional to the metabolic body size, or the 3/4 power of body weight, two animals may also be regarded as being on the same level of food intake, when they consume the same amount of metabolizable food energy per $kg.^{3/4}$

Not only the requirement of food energy, but also that of protein and of most vitamins, may be expressed per unit of the 3/4 power of body weight, because these dietary requirements are directly related to energy metabolism (see review by Kleiber 1945–46, p. 207).

For the dosage of drugs one should know whether or not the action depends on reaching a certain concentration in the blood stream without regard to its further maintenance. In this case the dosage should be proportional to body weight, since the amount of blood is proportional to body weight. If, however, the action of the biotic depends on the maintenance of a given concentration over a period of time, and if the rate of destruction or excretion of the biotic is proportional to the metabolic rate, then the dosage should be based on the metabolic body size.

The treatment may be so arranged that doses proportional to body weight are given and that the frequency of application depends on body size, so that over a sufficient period of time the rate of intake is proportional to the metabolic body size (the 3/4 power of body weight). In this case the frequency should be inversely proportional to the fourth root of body weight.[7]

If, for example, a 50 gram rat received one unit of a biotic daily, then to establish similar conditions under the assumptions made, a 500 kg. steer should receive a dose of 10,000 units every ten days.

BODY SIZE AND FOOD UTILIZATION. The establishment of a metabolic unit of body size is particularly advantageous for the investigation of food utilization.

$$\text{The quotient: } \frac{\text{energy in animal product}}{\text{total food energy}} = \frac{G}{U}$$

is called the total efficiency of food utilization. It measures (as far as that is possible in terms of energy) the success of animal husbandry.

[6] Metabolizable energy = heat of combustion of food minus heat of combustion of feces, urine, and methane.

[7] If q = amount given at one time (dosage)
f = frequency of application
W = body weight
then rate of intake $= f \cdot q = k_1 \cdot W^{3/4}$; $q = k_2 \cdot W$

$$f = \frac{k_1}{k_2} \cdot \frac{W^{3/4}}{W} = \text{const.} \cdot W^{-1/4}$$

The quotient $\frac{\Delta G}{\Delta U}$ is the partial efficiency. It is in general dependent on the level of food intake, but the discussion in this chapter may be limited to the simplest special case, that in which the partial efficiency is constant $= e$.

An animal needs a certain amount of food energy, U_m, to maintain itself without gain or loss of body substance. Only that amount of food energy consumed in excess of this maintenance requirement is available for production. The energy in the animal's production is, thus, the partial efficiency times the food energy available for production:

$$G = e(U - U_m)$$

therefore the total efficiency is:

$$\frac{G}{U} = e\left(1 - \frac{U_m}{U}\right)$$

The maintenance requirement U_m is related to the rate of fasting metabolism, B. The amount, B, of energy in body substance which the fasting animal would lose, is saved from being lost by the food energy U_m. The quotient $\frac{B}{U_m}$ is therefore the partial efficiency of food utilization for maintenance. Assuming for simplicity that this partial efficiency is the same as that for production, then $\frac{B}{U_m} = e$ or $U_m = \frac{B}{e}$.

Introducing this expression in the equation for total efficiency one obtains

$$\frac{G}{U} = e\left(1 - \frac{1B}{eU}\right), \quad \text{or } \frac{G}{U} = e - \frac{B}{U}$$

That means the total efficiency of food utilization is the difference between the partial efficiency (characterizing the nutritive content of the food), and the ratio between the basal metabolic rate and rate of intake of food energy. This ratio characterizes the capacity of the animal to take in food. Partial efficiency of utilization of metabolizable food energy (U) is related directly to the calorigenic (or specific dynamic) effect of food, ΔQ:

$$e = \frac{\Delta G}{\Delta U} = 1 - \frac{\Delta Q}{\Delta U}$$

There is no reason to assume that partial efficiency depends on body size. To the contrary, there is a good deal of evidence against such an assumption. Chambers and Lusk (1930) and Eaton, Cordill, and Gouaux (1936) for example observed that the specific dynamic action of glycine, administered to dogs varying from 5 to 13 kg., was independent of body size. The ratio $\frac{U}{B}$ has been defined as the "relative food level," and if an animal eats to capacity then $\frac{U}{B}$ may be termed the "relative food capacity" in terms of the basal metabolic rate.

Since the basal metabolic rate is proportional to the 3/4 power of body weight ($W^{3/4}$) the term $\frac{U}{W^{3/4}}$ characterizes the relative food capacity of the animal in terms of metabolic body size. Aside from differences in partial efficiency that animal is the better food utilizer which has the greater relative food capacity. The food intake per kg.$^{3/4}$, and similarly the gain per kg.$^{3/4}$, should be important characteristics for selection in breeding farm animals.

Since there is no reason to assume that partial efficiency depends on body size, the question of food utilization and body size is mainly a question of whether or not relative food capacity depends on body size. The simplest hypothesis is that the relative food capacity is independent of body size. Kleiber (1933), based on data obtained by himself and others, produced some empirical evidence in support of this simple hypothesis. By extensive experiments, Brody and his co-workers have confirmed this hypothesis for lactation (Brody and Proctor, 1935; Brody and Cunningham, 1936). Brody and Nesbit (1938) reported that energetic efficiency of rats (for lactation) is within that of cows. The efficiency for mechanical work is the same for large and small horses (Brody and Cunningham 1936) and for large and small men (Robinson, 1942).

A corollary to the hypothesis that total efficiency of food utilization is independent of body size, is the postulate that the relative rate of animal production is independent of body size. The relative rate of animal production may be the rate of gain in body substance per kg.$^{3/4}$, the rate of mechanical work per kg.$^{3/4}$ or the rate of milk production per kg.$^{3/4}$ (Kleiber and Mead, 1941, 1945).

Kleiber's hypothesis does not, of course, exclude the possibility that there are good and bad food utilizers. The meaning of the hypothesis is that the comparisons of good and large food utilizers with small and bad food utilizers, or vice versa, should not be used to establish a relation of body size and food utilization.

Similarly, one cannot directly compare young small animals with old large animals, even though in some cases age may not affect the relative production level. This is true, for example, for a considerable range in pigs. From figures given by Breirem (1939), one may calculate that the relative rate of gain for a 20 kg. pig is 112 kcal. per kg.$^{3/4}$; that of a 100 kg. pig, 118 kcal. per kg.$^{3/4}$.

SUMMARY AND CONCLUSIONS

1. Among homeotherms, from mice to cattle, metabolic rate and body size are correlated. This correlation is especially high when the metabolic rates are measured under standard conditions.

2. The metabolic rate of large and small homeotherms is more nearly proportional to the area of their respective body surfaces than to their body weights. This relationship is known as the surface law.

3. From five types of theories interpreting the surface law three are rejected, namely:

a. The theory based on the erroneous idea that the summated area of internal surfaces, such as the surfaces of the cells and the pulmonary alveoli, is proportional to the surface area of the body. The similarity in build of large and small animals cannot be extended to the dimensions of cells or alveoli.

b. Rejected also is the theory which bases the surface law on alleged differences in the chemical composition of large and small animals, involving such vague concepts as the "active protoplasmic mass". Neither does the theory which makes metabolic rate a function of anatomical composition stand the test. Uncritical evaluation of empirical data has led to the ill conceived generalization that the weights of the blood, the major glands and other organs are proportional to the 2/3 power of body weight or to the surface area. Logical application of this generalization can lead to absurd results.

c. The theory relating the surface law to an allegedly genetically fixed constant oxygen requirement of the cells, has rather little value for a physiological interpretation of the surface law. This theory, moreover, is in opposition to the well known fact that the metabolic rate of animals is essentially affected by somatic conditions. Furthermore, the metabolic rates of genetically closely related homeotherms of different size differ more than do the metabolic rates of genetically very different homeotherms but with the same body size. Admittedly, the concepts of "genetically similar or different" may change in the future when the biochemistry of the gene is further developed.

4. Two of the five types of theories on surface law are basically sound, namely, the theory connecting metabolic rate with rate of heat transfer and the theory relating metabolic rate to blood circulation. These theories may be integrated into one as follows: *In natural selection, those animals prove to be better fit whose rate of oxygen consumption is regulated so as to permit the more efficient temperature regulation as well as the more efficient transport of oxygen and nutrients.*

5. This theory does not postulate a strict proportionality between the area of a "true" body surface and metabolic rate. Body surface area has been used in very valuable work on animal heat exchange (Rubner, 1902; Deighton, 1933; Winslow et al., 1934–39; Burton, 1934; DuBois, 1937; Hardy and DuBois, 1938). As a basis for comparing metabolic rates of large and small animals, however, body surface area is not well enough defined. The use of an ever increasing variety of surface areas, even for the same species of animals, has led to an unnecessary and deplorable state of vagueness in comparative physiology of metabolism.

6. Relatively recent results on homeotherms, ranging from mice to cattle, indicate that the metabolic rate per unit of the surface area is greater the larger the animal. A linear correlation between the logarithm of metabolic rate and the logarithm of body weight shows that metabolic rate is proportional to a given power function of body weight. The metabolic rate divided by the 3/4 power of body weight is independent of body size.

7. The 3/4 power of body weight is therefore recommended as representative of metabolic body size, and "$kg.^{3/4}$" chosen as the symbol for the unit. The body weight in kilograms, raised to the 3/4 power, measures the metabolic body size of an animal in $kg.^{3/4}$.

8. The metabolic level of an animal may be characterized as the metabolic rate per $kg.^{3/4}$. Under standard conditions the metabolic level of adult homeotherms, from mice to cattle, averages 70 kcal. per $kg.^{3/4}$ per day or about 3 kcal per $kg.^{3/4}$ per hour.

9. At present there seems to be no sufficient reason against the intraspecific application of the 3/4 power rule of metabolic rate. Modulating factors for age and specific stature may be incorporated into prediction equations for human metabolic rate based on the 3/4 power rule. Such equations are preferable to the irrational regression formulas of Harris and Benedict, and they avoid the connection with the ill defined body surface involved in the Du Bois prediction.

10. The unit of metabolic body size is useful for expressing levels of food intake and of animal production, it is a sound basis for comparing food capacity and production capacity of animals that differ in body size. Relative food capacity (maximum rate of food intake per kg.$^{3/4}$) and relative production capacity (maximum rate of production per kg.$^{3/4}$) should be among the most important criteria for selecting efficient food utilizers. The metabolic body size may also be useful in estimating dosage of biotics.

11. When the concepts concerned with the relationship of body size and metabolic rate are clarified, and when not only the methods of measurement, but also those of reporting the data are sufficiently standardized, then comparative physiology of metabolism will be of great help in the efforts to solve one of the most interesting and intricate problems of biology, the regulation of the rate of cell metabolism.

Acknowledgment. I am grateful to A. H. Smith, graduate student in our department, for valuable assistance in the preparation of this review.

REFERENCES

Barrón, E. S. G. Advance Enzymol. **3**: 149, 1943.
Beer, E. J. de and A. M. Hjort. Am. J. Physiol. **124**: 517, 1938.
Benedict, F. G. Ergebn. Physiol. Exper. Pathol. **36**: 300, 1934.
The physiology of the elephant. Carnegie Inst. of Wash. Publ. **474**: 1936.
Vital energetics. Carnegie Inst. of Washington, Publ. **503**: 215 pp., 1938.
Benedict, F. G. and R. C. Lee. Annales de Physiol. **12**: 983, 1936.
Benedict, F. G. and E. G. Ritzman. Proc. Natl. Acad. Sci. **21**: 304, 1935.
Bergmann, C. and R. Leuckart. Anatomisch—physiologische Übersicht des Thierreichs. p. 217, Stuttgart, 1855.
Blank, H. Pflüger's Arch. **234**: 310, 1934.
Boothby, W. and I. Sandiford. Physiol. Rev. **4**: 69, 1924.
Breirem, K. Beretning Vet. Landbohojskoles **170**: 1936.
Tierernährung **11**: 500, 1939.
Brody, S. Bioenergetics and growth. Reinhold Publ. Co., New York, 1945.
Brody, S., J. E. Comfort and J. S. Mathews. Mo. Res. Bull. **115**: 1928.
Brody, S. and R. Cunningham. Mo. Res. Bull. **238**: 1936.
Mo. Res. Bull. **244**: 1936.
Brody, S. and R. Nisbet. Mo. Res. Bull. **285**: 1938.
Brody, S. and R. C. Procter. Mo. Res. Bull. **166**: 1932.
Mo. Res. Bull. **222**: 1935.
Brody, S., R. C. Proctor and U. S. Ashworth. Mo. Res. Bull. **220**: 1934.
Bruhn, J. M. and F. G. Benedict. Proc. Am. Acad. Art and Sci. **71**: 259, 1936.
Burton, A. C. J. Nutrition **7**: 497, 1934.
Chambers, W. and G. Lusk. J. Biol. Chem. **85**: 611, 1930.
Clark, A. J. Comparative physiology of the heart. Cambridge Univ. Press, 1927.
Deighton, T. Physiol. Rev. **13**: 427, 1933.
Donaldson, H. H. The rat. Philadelphia, 1924.

PHARMACOLOGY OF ANTIHISTAMINE COMPOUNDS

EARL R. LOEW

Department of Pharmacology and Experimental Therapeutics, University of Illinois College of Medicine, Chicago

Those drugs which are capable of diminishing or preventing several of the pharmacological effects of histamine and which do so by a mechanism other than the production of pharmacological responses diametrically opposed to those produced by histamine may be termed *histamine antagonists* or *antihistamine drugs*. Epinephrine and its congeners are not considered as antihistamine drugs, being excluded by definition since they induce prominent effects such as bronchiodilatation, vasoconstriction, decreased capillary permeability, inhibition of intestinal activity, etc., which represent the antithesis of those produced by histamine. True antihistamine drugs, on the other hand, are able to antagonize histamine without eliciting pharmacological responses, or if responses are elicited they do not appear to be of the type or degree which suggest an important causal relationship to the histamine antagonism. Excitatory sympathomimetic activity is not essential for antihistamine action since certain adrenergic blocking drugs are capable of antagonizing histamine. Use of the term "histaminolytic" with reference to drugs which antagonize histamine has been criticized (36) since there is no evidence that the drugs referred to are capable of destroying or neutralizing histamine.

In the present discussion, emphasis will be placed on those drugs which are potent histamine antagonists and which, from the evidence available, appear to act as blocking agents with some degree of specificity as regards antihistamine action. Pharmacological experience reveals that specificity is relative and not absolute, the latter being rare or practically unattainable. Atropine and other antispasmodic drugs, epinephrine and related amines, and certain amino acids and derivatives of histamine are capable of antagonizing some of the effects of histamine (for refs. cf. 78, 142, 150, 151). However, these compounds can be regarded as non-specific since their ability to contract or relax smooth muscle and to antagonize acetylcholine, barium and other spasmogenic agents equals or exceeds their effectiveness in opposing histamine.

The motivation for the development of antihistamine drugs undoubtedly came from the realization that histamine, omnipresent in tissues and able to produce diverse effects when minute quantities are liberated therefrom, possibly acts as a mediator or regulator of physiological processes and as a culprit in certain pathological states. Several reviews cite much of the literature dealing with the questions concerning the involvement of histamine in various physiological mechanisms (10, 43, 46, 166). Problems to be clarified include the rôle which histamine may play in mediating vasodilator nerve impulses (96), in altering capillary permeability (97, 102), in producing gastric secretion (51, 73, 87), pain (157), and reactive hyperemia (5).

Recent reviews discuss evidence related to the question as to whether histamine is one of the factors involved in producing certain manifestations in pathological conditions (10, 32, 43–46, 60, 149, 150, 166). Opinions concerning the rôle of histamine in various types of inflammatory reactions vary widely (13, 102, 131, 150, 152, 153, 159). There is an obvious need for determining the extent to which histamine may be involved in radiation effects (13, 114), certain types of headache (166, 188), gastric and duodenal ulcers, traumatic, burn and surgical shock (185), pruritus (25, 135, 159), responses to a variety of venoms (54, 94, 150), and toxemias of pregnancy. In the latter condition, and during normal pregnancy, there is the possibility that the metabolism of histamine and histidine may be appreciably altered (3, 92, 93). Prevalence of allergic diseases and the evidence which suggests that histamine accounts for some of the manifestations also emphasizes the importance of the development of antihistamine drugs.

In 1940 and 1941, the extensive critical reviews by Dragstedt (43, 44) and Feldberg (60) discussed the evidence which convincingly supported the hypothesis that histamine is liberated from tissues by the antigen-antibody reaction and accounts for the important symptoms of anaphylaxis. Likewise, in allergic diseases, a variety of alterations in tissues and fluids resemble the pharmacological responses to histamine and it is customary to consider the symptoms of allergy as resulting from an antigen-antibody reaction. Code (32), Rocha e Silva (149, 150) and Dragstedt (45) have recently presented additional evidence in support of the contention that histamine plays a prominent rôle in anaphylaxis and allergy.[1] These authors were cognizant of the fact that the evidence was indirect and circumstantial. All were aware of the need for additional evidence, even though indirect in type. The minute amounts of histamine which are capable of eliciting responses are probably not always detected by available chemical and biological methods.

Interest in antihistamine drugs is related to their use: *a*, as pharmacological tools for differentiating between effects of histamine and those due to other agents; *b*, as therapeutic agents to alleviate symptoms due to histamine, i.e., in allergy and possibly other conditions, and *c*, as diagnostic agents to aid in proving or disproving that symptomatology is referable to histamine or to other causes.

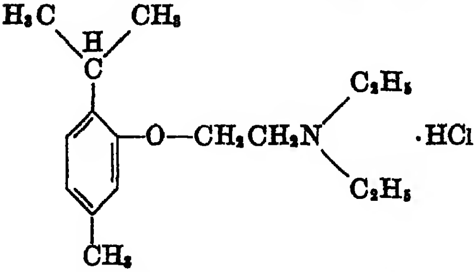
The use of antihistamine drugs in elucidating these numerous problems obviously necessitates an adequate knowledge of their pharmacological and toxicological properties, mode of action, degree of specificity, and in the absence of absolute specificity, a due consideration of other than antihistamine action.

The antihistamine drugs herein discussed will be identified before proceeding to a more detailed discussion of pharmacological action and toxicity.

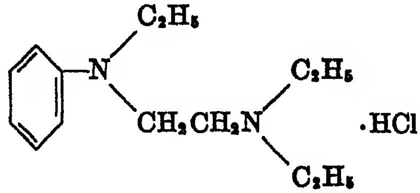
Fourneau compounds (929 F and 1571 F). 2-Isopropyl-5-methylphenoxyethyldiethylamine (929 F) is a phenolic ether which is frequently designated as thymoxyethyldiethylamine. Following the preliminary reports from Doctor

[1] The failure to detect an elevated blood level of histamine does not negate its involvement in tissue reactions since it leaves the blood stream with extreme rapidity (32, 47, 155) and traction methods do not always detect physiologically active amounts (53, 115).

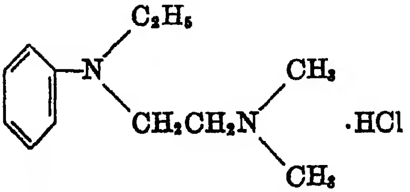
Bovet's laboratory[2] at the Pasteur Institute in Paris concerning the antihistamine and anti-anaphylactic properties of 929 F (23, 170), Anne-Marie Staub (169)



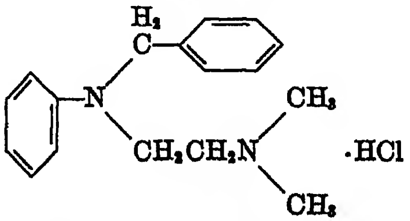
2-Isopropyl-5-methylphenoxy-ethyldiethylamine·HCl
(Fourneau 929)



N-Phenyl-N,N',N'-triethylethyl-enediamine·HCl
(Fourneau 1571)



N-Phenyl-N-ethyl-N',N'-dimethyl-ethylenediamine·HCl
(2325 Rhône-Poulenc)



N-Phenyl-N-benzyl-N',N'-dimethyl-ethylenediamine·HCl
(2339 Rhône-Poulenc) Antergan

published extensive data relative to the antihistamine properties of several series of compounds synthesized by E. Fourneau. 929 F proved to be the most effective of the phenoxyethylamines with respect to alleviation of the symptoms of histamine shock in guinea pigs, but was slightly less effective than 1571 F, N-phenyl-N,N',N'-triethylethylenediamine, selected from the series of N-phenylethylene diamines. The extensive pharmacological and toxicological studies conducted with the most promising antihistamine compounds, 929 F and 1571 F, deserve special commendation in view of the demonstration that agents could be synthesized which would antagonize histamine and also diminish or annul certain anaphylactic reactions. Furthermore, their findings were soon corroborated (24, 31, 77, 107, 110, 111, 112, 156) and rapidly extended (18, 27, 75, 76, 106, 132, 158, 178, 186, 187). The diverse pharmacological actions and toxic effects produced by 929 F and 1571 F have precluded their use in therapeutics and extensive use as research tools. The success of the French investigators motivated the search for superior antihistamine compounds.

Antergan (2339 R.P.) and 2325 R.P. Evidence of definite progress in the

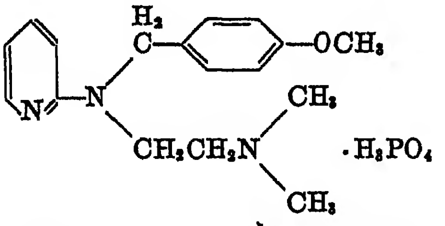
[2] The erroneous statement has been made (146) that the starting point for the develop-ment of antihistamine drugs began in 1933 with the report of Fourneau and Bovet (Arch. int. Pharm. et thérap., 46: 178–191, 1933). However, the phenolic ethers considered were investigated for sympatholytic properties and no reference was made to their effects in antagonizing histamine.

development of more potent antihistamine compounds was presented by Halpern (77) who studied twenty-four derivatives of 1571 F and selected two compounds which exerted marked antihistamine action from the compounds synthesized by Mosnier in the laboratories of Rhône-Poulenc. The dimethyl homologue of 1571 F, N-phenyl-N-ethyl-N′,N′-dimethylethylenediamine (2325 R.P.), and N-phenyl-N-benzyl-N′,N′-dimethylethylenediamine (2339 R.P. or Antergan) were selected and studied extensively relative to antihistamine, anti-anaphylactic, and antispasmodic action, as well as pharmacodynamic and toxicological properties. Of the aniline derivatives, 1571 F, 2325 R.P. and Antergan, the latter is the only compound which has been evaluated therapeutically in man (for ref. cf. 56, 146).

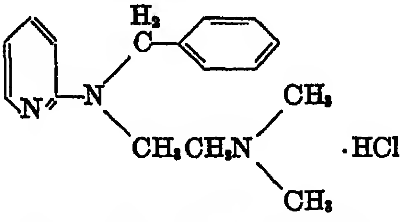
Antistine (2-N-phenyl-N-benzylaminomethylimidazoline), an analogue of Antergan, appears to possess pharmacological properties and exert therapeutic effects similar to those of Antergan (19, 25, 127, 162).

Since 1942, synthetic compounds other than aniline derivatives have been found to possess marked antihistamine properties and a sufficiently low toxicity as to permit effective use as research tools and therapeutic agents. The compounds which have been examined most extensively were selected from α-aminopyridine derivatives and benzhydryl alkamine ethers.

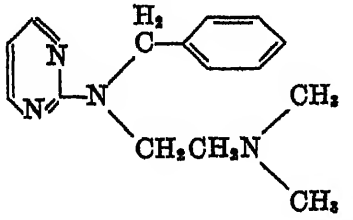
α-Aminopyridine derivatives. Definite progress was made by Horclois (cf. 24) who substituted a heterocyclic structure, pyridine, for the phenyl ring in Antergan. Bovet and co-workers (20, 21, 24), in 1944, found that N-α-pyridyl-N-p-



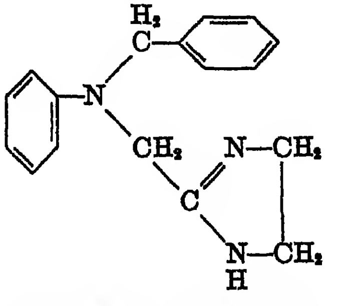
N-α-Pyridyl-N-p-methoxybenzyl-N′, N′-dimethylethylenediamine·H_3PO_4 (2786 R. P.) Neoantergan



N-α-Pyridyl-N-benzyl-N′,N′-dimethylethylenediamine·HCl Pyribenzamine hydrochloride



N-2-Pyrimidyl-N-benzyl-N′,N′-dimethylethylenediamine Hetramine



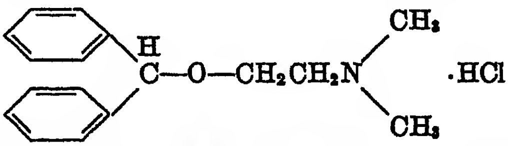
2-N-Phenyl-N-benzylamino-methylimidazoline Antistine

methoxybenzyl-N′,N′-dimethylethylenediamine (2786 R.P. or Neoantergan) possessed a remarkable degree of antihistamine potency, as shown by the relief or prevention of bronchospasm in guinea pigs induced by anaphylaxis or administration of histamine. This compound was found to be capable of diminishing or preventing most of the effects of histamine on smooth muscle and the vascular system.

In 1945, Mayer, Huttrer and Scholz (121, 122) reported the antihistamine and anti-anaphylactic action of a homologue of Neoantergan, N-α-pyridyl-N-benzyl-N′,N′-dimethylethylenediamine (63-C or Pyribenzamine hydrochloride). Pyribenzamine lacks the methoxy group on the benzyl ring contained in Neoantergan. The close chemical similarity between Neoantergan and Pyribenzamine would suggest similarities in the qualitative actions and toxicity of the two compounds which will be apparent from the discussion to follow.

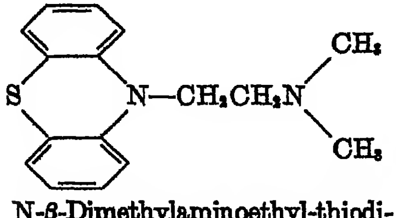
Hetramine, N-2-pyrimidyl-N-benzyl-N′,N′-dimethylethylenediamine, is a pyrimidine analogue of Pyribenzamine. In recent months, Feinstone, Williams and Rubin (59) demonstrated that Hetramine protected guinea pigs against histamine shock and anaphylactic shock.

Benzhydryl alkamine ethers. The availability of benzhydryl derivatives synthesized by Rieveschl and Huber (cf. ref. 110) which possessed an ether linkage and other structural characteristics similar to phenolic ethers such as 929 F, motivated the author, in 1943, to test the benzhydryl alkamine ethers for antihistamine action. β-Dimethylaminoethyl benzhydryl ether (Benadryl

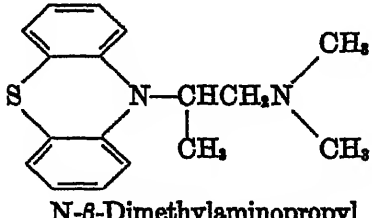


β-Dimethylaminoethyl benzhydryl ether·HCl
Benadryl hydrochloride

hydrochloride) and several related tertiary amines exerted definite antihistamine and anti-anaphylactic action (107, 110). Furthermore, it was demonstrated (109, 112, 187) that quaternary nitrogen derivatives of Benadryl were potent histamine antagonists and, although the quaternary compounds are choline derivatives, they also exerted an anti-acetylcholine or atropine-like action instead of the usual muscarinic action exerted by most choline ethers and esters (72.)



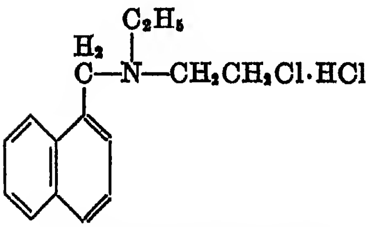
N-β-Dimethylaminoethyl-thiodiphenylamine
(3015 R. P.)



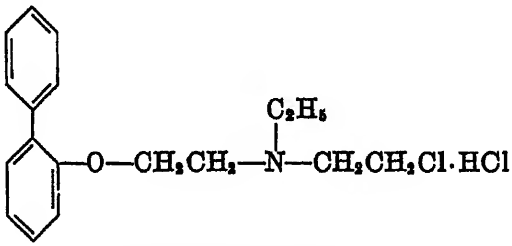
N-β-Dimethylaminopropyl thiodiphenylamine
(3277 R. P.)

Thiodiphenylamines. According to preliminary reports by Halpern and Ducrot (80) and Halpern (79), two derivatives of thiodiphenylamine (3015 R.P. and 3277 R.P.) protected guinea pigs, rabbits and dogs against exceedingly large doses of histamine injected intravenously. It was stated that these compounds antagonized the spasmogenic action of histamine on the uterus and intestine and the depressor action in the cat, dog, and rabbit. Salivary, pancreatic and gastric secretion (histamine-induced?) were not inhibited.

Other compounds. Of particular importance in relation to anti-anaphylactic effectiveness and mode of action of antihistamine drugs is the discovery (1) of a group of compounds which oppose several actions of histamine and, in contrast to all other potent antihistamine compounds which enhance the pressor action of epinephrine, these block and reverse the pressor action of epinephrine (adrenergic blocking action). α-Naphthylmethylethyl-β-chloroethylamine and β-2-biphenyloxyethyl-β-chloroethylethylamine[3] exert this dual action. They are effective in alleviating histamine shock and anaphylactic shock in guinea pigs and in diminishing the depressor responses to histamine in dogs. α-Naphthylmethylethyl-β-chloroethylamine and Neoantergan are equally effective (1) in preventing fatal histamine shock (M.E.D. of 0.025 mgm./kgm., subcutaneously) and anaphylactic shock in guinea pigs.



α-Naphthylmethylethyl-β-chloroethylamine·HCl



β-2-Biphenyloxyethyl-β-chloroethylethylamine·HCl

PHARMACODYNAMICS. Doses of antihistamine drugs commonly employed in animals do not exert a marked effect on respiration and the clinical literature contains no indication that doses recommended for therapeutic use in man have any effect (126, 146). An increased rate and amplitude of respiration persisted for some minutes following intravenous injection of 929 F, 1571 F (169) and Antergan (77) in dogs anesthetized with chloralose. Similar responses have been obtained with Neoantergan, Pyribenzamine and Benadryl (112, 167, 168, 191) in dogs anesthetized with pentobarbital or phenobarbital. A short period of apnea preceded the stimulatory phase when several of these drugs were injected rapidly (77, 112, 167). With all compounds excepting 929 F, it is doubtful that the respiratory alterations are reflex responses referable to changes in blood pressure since the brief apnea was usually concurrent with hypotension. Furthermore, respiratory stimulation persisted for some minutes

[3] Synthesised by Drs. G. Rieveschl, Jr., R. Fleming and W. R. Coleman of Parke, Davis and Company, Detroit, Michigan.

and, in general, was concurrent with a normal or slightly increased level of arterial blood pressure.

Following administration of large oral or parenteral doses of antihistamine drugs, respiratory stimulation has also been noted in unanesthetized animals (77, 105, 120). Such stimulation usually preceded or was concomitant with marked excitability, tremors and convulsions. Secondary respiratory depression may have been a contributory cause of death (120, 147). Mention was made of a general narcotic effect in rats which received large doses of Neoantergan (42).

The hypotension induced in anesthetized dogs, cats and rabbits with large doses of 929 F was frequently pronounced and of appreciable duration (31, 105, 169, 178), thus complicating interpretation of data relating to histamine antagonism. Hypertension was reported to occur in unanesthetized dogs (132) treated with 929 F (40 mgm./kgm., S.C.). More potent antihistamine drugs such as 1571 F, Antergan, Neoantergan, Benadryl and Pyribenzamine, induced a transient hypotension in dogs when injected intravenously at a rapid rate (77, 112, 167, 168, 169, 190). However, it is quite apparent that hypotension was minor in degree, absent, or replaced by a slight hypertension which persisted for several minutes when these drugs were administered slowly or by other routes (24, 77, 112, 137, 167, 169, 182, 191). The exact cause of the hypotension or hypertension has not been determined although hypertension was considered to be in part referable to a stimulation of the central nervous system by Antergan (77) and Pyribenzamine (191). Following administration of Antergan, vasoconstriction occurred in the intestine, spleen and kidney, whereas vasodilatation was induced in extremities (40, 77). DeCuyper (40) reported that hypertension induced with Antergan was not dependent on sino-carotid reflexes or liberation of epinephrine from the adrenal glands, and concluded that it was due to a direct, peripheral vasoconstrictive action.

The rise in arterial pressure in anesthetized dogs following intravenous injection of Antergan, Pyribenzamine, Neoantergan and Benadryl was slight in degree. Since the hypertension persisted for only a few minutes, it could scarcely be related to other actions of the drugs, such as antagonism of histamine, which was demonstrable for one or more hours. Doses of antihistamine drugs which are therapeutically effective in allergic diseases do not raise blood pressure (70, 126, 146). This indicates that antihistamine drugs lack the pressor properties possessed by epinephrine and its congeners.

The majority of antihistamine drugs exert some local anesthetic action (25, 31, 42, 77, 110, 157, 169), the degree of which has not been quantitatively determined except with Neoantergan (42). Local anesthetic action should be considered when interpreting effects of antihistamine drugs on vascular and cutaneous responses in which axon and other reflexes are involved. Such action could possibly be related to anti-pruritic effects and production of cutaneous analgesia (31, 42, 157, 158), the latter having been demonstrated following large doses of 929 F and Neoantergan in animals.

Very few studies have been made which concern the direct and indirect action of antihistamine drugs upon the heart. Neoantergan was reported to exert a

quinidine-like action on isolated auricular tissue from rabbits (42). In normal human subjects, oral administration of maximal therapeutic doses of Benadryl for several weeks seldom lowered systolic blood pressure and did not produce clinically detectable alterations of the heart (126).

Toxic doses of all antihistamine drugs, with the exception of quaternary derivatives of Benadryl (187), stimulate the central nervous system of animals as evidenced by hyperexcitability, tremors and convulsions (cf. toxic effects). It has been stated (29) that Antergan enhances the excitability of cortical fibers and peripheral motor and sensory fibers; these effects are in accord with the generalized excitation produced in animals by Antergan (77). In animals, antihistamine drugs do not induce sedation, hypnosis or depressant effects. In contrast to the hyperexcitability and convulsions induced in several species of animals with large doses of antihistamine drugs is the sedation and dizziness which sometimes obtain in human beings following administration of therapeutic doses. In man, sedation is unpredictable, variable in degree from person to person, and constitutes a side-effect which has been encountered frequently with Benadryl (9, 55, 58, 101, 123, 135, 146, 172), less frequently with Pyribenzamine (6, 7, 58, 63, 146) and which is known to occur following use of Antergan (37, 39) and Neoantergan (37, 38).

Untoward reactions in man which result from direct or indirect actions of antihistamine drugs on the gastrointestinal tract include gastric distress, nausea, emesis, colic and diarrhea. Such reactions are seldom encountered with Benadryl (9, 55, 101, 123, 135, 172) which exerts a slight degree of antispasmodic action (vide infra). They possibly occur more frequently with Antergan (28, 37, 163, 164), Neoantergan (37) and Pyribenzamine (6, 7, 58, 63, 146), i.e., with those drugs which are practically devoid of antispasmodic properties, but which are capable of inducing spasm of intestinal and uterine muscle of animals *in vitro* and *in vivo* (vide infra).

Antihistamine drugs are absorbed rapidly following oral and parenteral injection. The usual experimental and therapeutic doses exert antihistamine and anti-allergic action in animals and man for two to six hours. No information is available concerning effective blood levels and rate or mode of destruction or elimination. Analyses made by Golvin and McGavack (71) revealed the presence of comparable quantities of Benadryl in the blood and spinal fluid of seven patients.

Acute and chronic toxicity. Those antihistamine drugs which are now being used as therapeutic agents were developed concurrently, or in rapid succession, which fact accounts for the minimal number of direct comparative studies relating to pharmacodynamics and toxicity. Comparisons of toxicity data obtained in different laboratories under varying conditions relative to technique and animals employed, may be misleading, especially when no reference compound is included in the studies which would at least permit a comparison of relative toxicity values. A direct comparison (167) of the acute toxicity following intraperitoneal injection of Neoantergan, Pyribenzamine and Benadryl in mice revealed no important difference in the values for LD 50 which ranged

from 75 to 90 mgm./kgm. Acute toxicity data (LD 50) reported by various investigators are included in table 1.

The tabulated data reveal that none of the antihistamine drugs are extremely toxic and it is unlikely that any major differences exist in relation to the acute toxicity of the various compounds. Furthermore, the ratio of these lethal doses to the small doses required to demonstrate antihistamine and anti-anaphylactic activity in animals consistently yields a high therapeutic index. The magnitude of the toxic doses of 929 F, 1571 F (31, 110, 169), and 2325 R.P. (24, 77) in

TABLE 1
Acute toxicity of antihistamine drugs (LD 50 in mgm./kgm.)

ANIMAL AND ROUTE OF INJECTION	ANTERGAN		NEOANTERGAN		BENADRYL		PYRIBENZAMINE	
	LD 50	(Ref. no.)	LD 50	(Ref. no.)	LD 50	(Ref. no.)	LD 50	(Ref. no.)
Mouse:								
i.p.			90	(167)	75	(167)	80	(167)
s.c.	175	(24)	150	(24)	130	(147)	75	(120)
oral					167	(147)	210	(120)
i.v.	40	(24)	30	(24)			12	(120)
Rat:								
i.p.					82	(110)		
s.c.	175	(77)					340 ♂ 225 ♀	(120)
oral	300	(77)			545	(147)	570 ♂ 515 ♀	(120)
i.v.					46	(147)	12	(120)
Guinea pig:								
s.c.	110	(24)	70	(24)				
i.p.					75	(110)		
Rabbit:								
s.c.							33	(120)
i.v.					10.5	(147)	9	(120)
Dog.:								
i.v.					30	(147)		

several animal species at least suggests that these drugs are no more acutely toxic in animals than compounds included in table 1, but 929 F produced marked side actions in dog and man (31, 169).

Toxic doses of antihistamine drugs in several animal species induced tremors, then convulsions which were sometimes followed by depression, and finally death which appeared to be due to cardio-respiratory depression. The exact cause of death should be determined, as well as means of antidoting toxic amounts of these drugs. The acute toxic effects of antihistamine drugs appear unrelated to any physiological mechanism involving histamine.

No indications of chronic toxicity were observed in studies on blood, urine, parenchymatous organs, nervous system or bone marrow of dogs receiving Benadryl orally (147) for several weeks in daily doses (40 to 60 mgm./kgm.) capable of causing nervousness, ataxia, gastro-intestinal reactions and hyperesthesia of the skin. In man, administration of Benadryl for as long as eighteen weeks did not cause variations in cells and other constituents of blood (126). Pyribenzamine has been administered orally to three human subjects for 80 days (95), to four dogs for one year and to rats for many months (117, 120) without inducing biochemical or pathological changes in body fluids, urine, or tissues. Doses of Neoantergan equivalent to the maximum human therapeutic dose, on the basis of body weight, were given to young rats in two injections daily for eleven days without affecting growth (42).

It is encouraging to note that the clinical literature referable to Antergan, Benadryl, Pyribenzamine, Neoantergan and Antistine yields no evidence as yet that they can induce blood dyscrasias or organic changes in the nervous system or parenchymatous organs.

ANTAGONISM OF HISTAMINE. A prominent pharmacological action of histamine is its spasmogenic action on smooth muscle of bronchioles, gastro-intestinal tract, uterus, ureters and gall bladder (10, 61, 72). Histamine induces arteriolar constriction or dilatation depending on animal species or locality of vascular bed in a given species. A conspicuous action of histamine is the production of increased capillary dilatation and greater capillary permeability, the latter factor being most important in the production of localized edema (97). Finally, histamine is a potent secretagogue which induces secretion from lacrimal, salivary, gastric and pancreatic glands. This secretagogue action is probably most important with reference to gastric secretion.

A consideration of antihistamine drugs must necessarily concern their ability to antagonize the various actions of histamine.

1. *Bronchioles.* Usually, antihistamine drugs have been selected and partially evaluated by determining their ability to alleviate bronchioconstriction which is so conspicuous in guinea pigs when histamine is injected or liberated during anaphylaxis (cf. 183). Until recently, it was difficult to compare the relative effectiveness of various drugs since various investigators used different techniques and frequently failed to include a standard or reference compound which would at least permit indirect comparisons of data from various sources. More recently, a majority of the antihistamine drugs were tested (65, 108, 110, 111, 167, 187) under conditions which were sufficiently standardized as to yield a remarkable consistency in the incidence of asphyxial deaths referable to histamine-induced bronchioconstriction in untreated guinea pigs, and which permitted a fairly reliable determination of the minimal effective dose (M.E.D.) of a given drug which was capable of partially annulling the action of histamine. Essentially, the method (110), a modification of that used by Kallós and Pagel (91) and others (77, 161), consists of determining the M.E.D. of a drug which significantly reduced the mortality rate of guinea pigs exposed to a lethal dose of an atomized histamine solution (see table 2).

The ability of 929 F and 1571 F to diminish the bronchioconstrictive effects of histamine in guinea pigs (23, 169, 170) has been confirmed by several investigators (31, 77, 110, 111, 156, 186). Antergan (77, 78) and 2325 R.P. were more effective. Antergan exerted a high degree of antihistamine specificity on the bronchiolar muscle of guinea pigs since even large doses failed to relieve bronchioconstriction induced with an aerosol of acetylcholine and escrine (77).

TABLE 2

Minimum effective doses of antihistamine drugs which reduce the incidence of mortality in guinea pigs exposed to an aerosol of histamine

DRUG	MOL. WT.	DOSE *	ROUTE	REF. NO.
		mgm./kgm.		
2-Isopropyl-5-methylphenoxyethyldiethylamine hydrochloride, 929 F (Fourneau)	285.80	6.0 6.0	i.p. s.c.	(110, 111) (24)
N-Phenyl-N,N′,N′-triethylethylenediamine hydrochloride, 1571 F (Fourneau)	256.81	3.0 5.0	i.p. s.c.	(110, 111) (77)
N-Phenyl-N-ethyl-N′,N′-dimethylethylenediamine hydrochloride, 2325 R.P. (Rhône-Poulenc)	228.75	2.5	s.c.	(77)
N-Phenyl-N-benzyl-N′,N′-dimethylethylenediamine hydrochloride, 2339 R.P., Antergan	290.81	1.0 0.5	s.c. s.c.	(24) (77)
N-Pyridyl-N-p-methoxybenzyl-N′,N′-dimethylethylenediamine·H_3PO_4, 2786 R.P., Neoantergan	382.37	0.075 0.025 0.1	i.p. s.c. s.c.	(167) (1) (24)
N-α-Pyridyl-N-benzyl-N′,N′-dimethylethylenediamine hydrochloride, Pyribenzamine	291.79	0.3	i.p.	(167)
β-Dimethylaminoethyl benzhydryl ether hydrochloride, Benadryl	291.45	1.5	i.p.	(110, 111,) (112, 167)
β-Benzhydryloxyethyltrimethylammonium iodide	397.30	0.5	i.p.	(112, 187)
N-2-Pyrimidyl-N-benzyl-N′,N′-dimethylethylenediamine, Hetramine	255.33	0.3 to 0.6	i.p.	(59)
α-Naphthylmethylethyl-β-chloroethylamine hydrochloride	284.19	0.025	s.c.	(1)

* Doses employed by Halpern (ref. 77) were capable of preventing development of serious symptoms, whereas those employed by other workers diminished mortality. Pretreatment with drugs was instituted 15 to 30 minutes before exposure to histamine aerosol.

Large doses of epinephrine were required to oppose effects of either acetylcholine or histamine.

In rapid succession, it was demonstrated that Neoantergan (21, 24), Benadryl and related compounds (109–112, 167), Pyribenzamine (118, 122), Antistine (127) and Hetramine (59) were highly effective in antagonizing the bronchioconstriction induced in intact guinea pigs by exposure to a histamine aerosol. Most of these findings have been confirmed (65, 66, 105, 167), and each of these drugs was capable of antagonizing the bronchioconstriction induced by intravenous injections of histamine.

The relative potency of several antihistamine drugs has also been ascertained by determining the number of intravenously injected LD 50 doses of histamine which a fixed dose of drug will antagonize sufficiently to prevent asphyxial death in guinea pigs (24, 66, 80). In general, the same relative potency of the various compounds is obtained by administering histamine intravenously or as an aerosol. Those drugs which exert a moderate degree of antihistamine activity are: 929 F, 1571 F, 2325 R.P., Antergan, Benadryl and several quaternary nitrogen derivatives, Pyribenzamine, Hetramine, Antistine and several alkyl derivatives of β-2-biphenyloxyethyl-β-chloroethylamine (1). A remarkable degree of potency was exhibited by Neoantergan, alkamine derivatives of thiodiphenylamine (3015 R.P. and 3277 R.P.) and α-naphthylmethylethyl-β-chloroethylamine (1). 3277 R.P. prevented immediate deaths following intravenous injection of 1500 lethal doses of histamine (80) in guinea pigs, whereas epinephrine and Aleudrine were incapable of protecting against two lethal doses.

It is known that the relative potency of drugs based on reactions in a given organ or tissue may be decidedly different when based on reactions in other tissues or organs. Because of this fact it is highly important to ascertain the degree of effectiveness of antihistamine drugs with respect to the antagonism of the action of histamine on the vascular system, intestinal and uterine smooth muscle, etc. Unfortunately, there is a paucity of data referable to such comparisons (vide infra), which is regrettable since a demonstrable degree of selective activity in certain tissues is important when selecting drugs for experimental and clinical use.

With several antihistamine drugs it has been shown that histamine-induced bronchioconstriction was diminished in isolated, perfused, guinea-pig lungs (49, 169, 192) and a limited amount of evidence indicated diminution of the histamine effect in other animals (50, 192) and man (35). Hewitt and Curry (84) found that Benadryl ameliorated bronchioconstriction and systemic reactions to impure preparations of streptomycin which contain histamine or a similar substance.

2. *Intestinal muscle.* Staub (169) reported that 929 F and 1571 F in a dilution of 2×10^6 to 1×10^7 prevented histamine from contracting isolated guinea-pig ileum. There is agreement (31, 77, 112, 187) that sufficient doses of 929 F or 1571 F suppressed the spasmogenic action of acetylcholine and barium, thus indicating that these antihistamine compounds are not absolutely specific as was suggested for 929 F by Rosenthal and Brown (156). According to Halpern (77), the dilution of drugs required to completely suppress the spasmogenic action of histamine (3.3×10^6; base?) on guinea pig intestine was 6.6×10^4 for 1571 F, 6.6×10^5 for 2325 R.P. and 1×10^7 for Antergan. In order to antagonize the spasmogenic action of acetylcholine and barium it was necessary to increase the concentration of Antergan and 2325 R.P. several hundred times, the latter compound being more specific but one-fifteenth as potent as Antergan in antihistamine action.

Other direct comparisons of the antihistamine action of several drugs on intestinal muscle were made by Loew *et al.* (112) who found that spasm induced

with histamine diphosphate (1.25×10^7) was reduced 75 to 100 per cent by 1571 F in a dilution of 3.3×10^6 and by Benadryl in a dilution of 5×10^7, whereas antispasmodics such as atropine, Pavatrine, and Trasentin were only effective in lower dilutions of approximately 1×10^6. It was also revealed that Benadryl, and especially 1571 F, were less effective than Pavatrine and Trasentin as regards antagonism of the spasmogenic action of barium and acetylcholine. Thus, the most pronounced action of 1571 F and Benadryl was their antagonism to histamine, Benadryl being most potent and 1571 F the most specific. An extension of these studies (109, 187) with several antispasmodics, Benadryl and some of its quaternary derivatives, revealed interesting differences in potency and specificity relating to antagonism of histamine, barium and acetylcholine. Histamine-induced intestinal spasm was diminished by Antistine (127), a drug which exhibited a moderate degree of specificity.

Since Neoantergan is markedly effective in relieving histamine-induced bronchioconstriction, it is unfortunate that there are no quantitative and comparative data regarding its antihistamine action with reference to isolated intestinal muscle. It has been stated (42) that Benadryl was much less effective than Neoantergan and that the inhibiting dose of Neoantergan was less than the dose of histamine required to induce spasm (24). Mayer (118) stated that a dilution of Pyribenzamine of 5×10^7 was sufficient to prevent the spasmogenic action of histamine in a dilution of 1×10^7 and occasionally 1×10^6. Pyribenzamine (118) and Neoantergan (42) weakly antagonized the spasmogenic action of acetylcholine.

In anesthetized dogs, spasm of the intestine induced with histamine has been blocked by 929 F (169), Pyribenzamine, Neoantergan and Benadryl (167).

The antispasmodic action of antihistamine drugs as evidenced by their ability to diminish the tonus and spontaneous motility of isolated rabbit intestine, or the intestine of intact dogs, was not marked in the case of 929 F and 1571 F (76, 169), 2325 R.P. (77), Neoantergan and Pyribenzamine (42, 167). In fact, Neoantergan and Pyribenzamine stimulated intestinal and uterine activity when injected intravenously in anesthetized dogs, whereas Benadryl decreased tonus and motility of the intestine (167). In unanesthetized dogs, 929 F usually had only a variable and transient inhibitory action on the stomach, duodenum and jejunum and no important effect on the terminal part of the ileum (76).

In view of these findings it is quite apparent that the majority of antihistamine drugs do not prevent spasm or exert a prominent relaxing effect on intestinal muscle except under conditions in which histamine has produced increased tonus, hypermotility or spasm. Benadryl exerts a weak antispasmodic action.

3. *Uterine muscle.* Histamine has a powerful oxytocic action. Mayer (118) stated that 929 F and 1571 F were capable of contracting the isolated guinea-pig uterus, their spasmogenic action being sufficient to interfere with a demonstration of antihistamine activity. However, Staub (169) had previously reported that 929 F, in a dilution of 2×10^5, contracted the uterus whereas a dilution of 1×10^5 failed to do so and antagonized the oxytocic action of histamine. Mayer (118) stated that the antihistamine activity of Antergan and

Pyribenzamine was sufficiently higher than their spasmogenic activity to permit a demonstration of histamine antagonism. Halpern (77) reported that both Antergan and 2325 R.P. exerted a strong excitant action on isolated guinea-pig uterus. According to Bovet and Walthert (24), both Antergan and Neoantergan exhibited a degree of spasmogenic action which precluded a satisfactory demonstration of specific antihistamine action on the isolated guinea-pig uterus, although Dews and Graham (42) demonstrated histamine antagonism with Neoantergan in a dilution of 5×10^8; with rat uterus, the inhibiting action of histamine was not altered by non-spasmogenic amounts of Neoantergan. An obvious need exists for quantitative data obtained from several animal species to replace the statements, opinions and "typical" experiments which concern the spasmogenic action of 929 F, 1571 F, Antergan, 2325 R.P., Pyribenzamine and Neoantergan, and the ability of these drugs to antagonize the oxytocic action of histamine. No data are available concerning the action of Benadryl on the isolated uterus.

In anesthetized dogs, the intravenous injection of Pyribenzamine and Neoantegan in doses of 3.0 mgm. per kgm. caused contraction of the uterus, whereas the same doses of Benadryl failed to do so (167); each of the drugs antagonized the contracting effect of injected histamine.

The limited data available warrant the tentative conclusion that the antihistamine agents, with the exception of Benadryl, are capable of contracting uterine muscle, and that all antagonize the spasmogenic action of histamine. Although Benadryl exerted a weak antispasmodic action on bronchiolar and intestinal muscle, none has yet been experimentally demonstrated for uterine muscle. The reviewer is unaware of any evidence whatsoever concerning the ability of antihistamine drugs to oppose actions of pituitrin on the uterus or other tissues. Quantitative data relative to effects of antihistamine drugs on uterine tissue of several species of animals are needed, and studies should determine the effectiveness and specificity with regard to antagonism of various spasmogenic agents, especially pituitrin.

4. *Arterial blood pressure.* Although Staub (169) reported that 929 F failed to antagonize the depressor action of histamine in chloralosed dogs, positive results were later demonstrated (31, 132) when small doses of histamine were employed. In rabbits, histamine shock (178) and hypertension induced with small doses of histamine have been diminished by 929 F (169) and hypotension induced with large doses of histamine in unanesthetized rabbits was diminished by 1571 F, Antergan and Neoantergan, the latter drug being effective in intravenous doses of 0.5 to 1.0 mgm. per kgm. of body weight (143).

Reports by Bovet and co-workers (20, 21, 24) revealed that Neoantergan was markedly effective in diminishing the depressor action of histamine in anesthetized dogs; nasal plethysmograph measurements indicated lessened vasodilatation. Pyribenzamine (167, 189) and Benadryl (112, 167, 181, 182) also diminished the depressor action of histamine in dogs. Comparative studies with Neoantergan, Pyribenzamine and Benadryl in dogs anesthetized with pentobarbital (167) revealed that each drug in intravenous doses of 3.0 mgm. per

kgm. body weight diminished the depressor action of histamine the same degree (*ca.* 55 per cent). Thus, these drugs appeared to possess equal potency with regard to antagonizing the depressor effects of identical doses of histamine in dogs even though their ability to antagonize the bronchioconstrictive action of histamine in guinea pigs varied widely. However, examination of data relating to Benadryl-histamine antagonism (112, 182) reveals that the percentage inhibition of the depressor action of histamine only varied within small limits when the dose of Benadryl was varied several-fold. It is therefore possible that quantitative data derived from experiments of suitable design may reveal differences in potency of these drugs. Neoantergan and thiodiphenylamine derivatives (3015 R.P. and 3277 R.P.) prevented death of dogs, rabbits and guinea pigs following tremendous doses of histamine injected intravenously (80).

Benadryl produced some diminution in the depressor response to acetylcholine (112, 167, 168) and since it weakly antagonized the spasmogenic action of acetylcholine on intestinal muscle (108, 109, 112, 187) and was capable of producing mydriasis and altering accommodation (82, 109, 126), it is apparent that Benadryl has a weak atropine-like action. The evidence available indicates that 1571 F, Antergan, 2325 R.P., Neoantergan and Pyribenzamine have only a slight atropine-like action (24, 77, 169) on intestines and bronchioles and the latter two compounds fail to diminish the depressor action of acetylcholine (167).

The essential pharmacological difference between Benadryl and the α-aminopyridine derivatives appears to be that Benadryl is less specific as an antihistamine drug by virtue of a higher degree of anti-acetylcholine action which may be expressed as an antispasmodic effect not appreciably apparent with Neoantergan and Pyribenzamine. In fact, the latter are spasmogenic in certain concentrations. Quantitatively, Neoantergan and Pyribenzamine were more effective than Benadryl in diminishing severity of histamine shock in guinea pigs (65, 66, 167) but the data now available are not sufficient to prove that there is a wide difference in effectiveness in antagonizing actions of histamine on the vascular system. Since the clinical literature reveals very close similarity in the ability and inability of Benadryl and Pyribenzamine to alleviate symptoms of various allergic diseases, it is apparent that clinical effectiveness is not directly related to degree of antihistamine action in guinea pigs.

It may be concluded that antihistamine drugs not only antagonize excitatory effects of histamine on bronchiolar, intestinal and uterine smooth muscle, but also prevent the relaxing or inhibitory effects of histamine on the vascular system, as indicated by antagonism of the depressor action of histamine in dogs and its coronary dilating action (42) in cats.

The data now available relative to the effects of antihistamine drugs on bronchiolar, uterine and intestinal smooth muscle, indicate that the term musculotropic antispasmodic fails to characterize definitely drugs which are potent and appreciably specific as antagonists of the spasmogenic action of histamine. The term antispasmodic would not be applicable with reference to the ability of these drugs to prevent histamine from *relaxing* smooth muscle of systemic arterioles and coronary vessels in carnivores, or from increasing permeability

of capillary endothelium (vide infra). For these reasons it becomes apparent that the term "antihistamine" is most useful in describing and classifying these drugs.

5. *Capillary permeability and cutaneous reactions.* The question regarding the rôle which histamine plays in inflammatory reactions is important for such reactions are elicited by thermal, chemical, mechanical and photic stimuli; those reactions which are referable to bacteria or allergens are intimately related to immunological processes. In the localized inflammatory response, the flare and wheal are related, respectively, to vasodilatation and accumulation of edema fluid. The immediate cause of vasodilatation and especially of localized edema in inflammatory reactions is therefore of importance and increased capillary permeability is an important factor relative to the transudation of fluid from capillaries to injured tissues (97, 131).

Of the substances known to increase capillary permeability, histamine is probably the most potent (34, 102) and comparatively large amounts are present in skin (4, 10, 74, 116, 155, 157). Histamine appears to be somewhat specific in inducing cutaneous reactions for it was effective at a concentration of M/10,000 whereas a variety of other amines, imidazoles and miscellaneous substances were found to be effective only in concentrations of M/2 to M/100 (34).

The measurement of cutaneous responses to histamine, allergens, and other substances are semi-quantitative and usually restricted to observations, or actual measurement, of the diameter of the area showing a flare and wheal. The effects of antihistamine drugs on these cutaneous reactions have seldom been measured in a precise manner and the scope of the various studies has been limited. For these reasons, the discussion regarding the effect of antihistamine drugs on localized cutaneous reactions to a variety of stimuli is largely descriptive rather than explanatory.

In the rabbit and dog, extravasation of dye has been used as a measure of increased capillary permeability. Benadryl (98), Neoantergan (14, 15, 98) and Pyribenzamine (145) diminished or annulled the localized accumulation of dye which usually follows intradermal injection of histamine. Benadryl did not protect the cutaneous capillaries of the rabbit against other injected or liberated substances (98) since trypan blue responses to trypsin, venom of Crotalus adamanteus, staphylococcus toxin (Burky strain Ha), heparin, tetracaine (Pontocaine), codeine, and horse serum were unaltered in non-sensitized rabbits. Although Friedlaender and Feinberg (64) believed that local application of Benadryl reduced a wheal and flare response to codeine sulfate (scratch test) in man it is pertinent to consider that a 5 per cent solution of Benadryl would exert local anesthetic action and thus affect axon reflexes concerned in flare reactions. Data from sensitized rabbits were suggestive but did not prove or disprove that histamine was importantly involved in the cutaneous response to antigen (98). None of the evidence obtained or cited suggested that choline or an ester was concerned in the responses to any of the agents employed.

The experimental data obtained from rabbits (98) strongly suggest a, that the action of antihistamine drugs on capillary permeability was limited to an

antagonism of histamine, and b, that trypsin, snake venom, and staphylococcus toxin increased capillary permeability through some mechanism which did not involve histamine. The studies made with non-sensitized rabbits provided no support for the contention that intradermal injections of various agents causes liberation of histamine which accounts for increased capillary permeability (154), although similar studies with other antihistamine drugs should be made in several animal species.

The wheal and flare referable to intracutaneous injections of histamine in man were frequently diminished following oral administration of Antergan (28, 136, 139, 176, 179), Neoantergan (176), Benadryl (48, 57, 64, 126, 141) and Pyribenzamine (7). The highest incidence and degree of diminution of cutaneous responses to histamine occurred when threshold doses of histamine were injected in subjects treated with adequate doses of antihistamine drugs.

This evidence that antihistamine drugs diminish cutaneous reactions to administered histamine suggests that the drugs should diminish reactions to substances known or suspected of containing histamine. Trypan blue responses of equal intensity produced in dogs by histamine and ascitic fluid were attenuated proportionally by Neoantergan (15), thus suggesting that ascitic fluid contains histamine. Other evidence of the presence of histamine in ascitic fluid and pleural exudates has been adduced (12, 15, 152, 153). It is conceivable that some of the reactions observed by Menkin (129–131) in his studies on inflammation which were attributed to a specific factor, leucotaxine, may have been due to histamine.

More data are needed relative to the antagonism of the cutaneous effects of snake (136, 139) and bee venoms (2). It was stated that Antergan therapy had no effect on erythema induced by ultraviolet irradiation of a limited area of human skin (139).

Frequent attempts have been made to diminish localized anaphylactic reactions in the skin of animals and man by pretreatment with antihistamine drugs. Large doses of 929 F and Antergan in tuberculous guinea pigs did not minimize the severe cutaneous reactions produced by injection of 10 mgm. of crude tuberculin intradermally; Antergan failed to diminish cutaneous responses to smaller amounts of tuberculin (16). In man, the cutaneous responses to both histamine and tuberculin were diminished by Antergan (26).

Hemorrhagic cutaneous reactions (Schwartzman), which followed intravenous injections of bacterial toxin in rabbits previously sensitized locally, were diminished by Neoantergan administered subcutaneously in divided doses over a period of ten hours (24). Failure of 929 F and Antergan to diminish such hemorrhagic responses (17) in several rabbits could have been referable to the use of large doses of toxin and inadequate duration of treatment with drugs less potent than Neoantergan.

Support for the concept that liberation of histamine from tissues accounts in part for the allergic type of inflammatory reaction is the evidence that antihistamine drugs diminish dermographic responses (8, 57, 64, 136, 139, 163) in man and the responses induced by intradermal injection of allergens (7, 64,

176). In addition, the clinical literature reveals that a comparatively high percentage of patients with acute or chronic urticaria derive definite symptomatic relief when treated with antihistamine drugs. Diagnostic skin-testing procedures effected to determine the presence of hypersensitivity should obviously not be conducted in patients who have recently ingested an antihistamine drug.

Antihistamine drugs probably diminish whealing by preventing histamine from increasing capillary permeability or by acting directly to decrease permeability. If the action is direct, antihistamine drugs should diminish increased capillary permeability caused by a variety of agents which do not contain or liberate histamine. According to Gelvin, Elias and McGavack (70), Benadryl did not affect the permeability of meningeal capillaries which are permeable to Benadryl since comparable quantities of the drug were found in blood and spinal fluid of seven patients (71). According to the prevailing concept, the erythematous flare induced by cutaneous injury depends upon axon reflexes. Evidence obtained by Parrot and Lefebvre (138) indicated that Antergan prevented histamine from initiating such reflexes; the failure of Antergan to prevent vasodilatation in dogs following electrical stimulation of antidromic nerve fibers rendered it unlikely that histamine was acting as a mediator.

It will be important to determine the extent to which the action of antihistamine drugs is referable to a direct effect upon capillaries, an antagonism of the effect of histamine on permeability and ability to initiate or mediate axon reflexes, or an enhancement of the effects of epinephrine (cf. discussion concerning specificity). Antihistamine drugs which fail to enhance, or actually block, effects of epinephrine may prove useful in elucidating the effects upon capillaries.

Since antihistamine drugs were capable of diminishing the vasodepressor action of histamine and its action of increasing capillary permeability it is significant to note that Benadryl (85) and Antergan (88) failed to induce beneficial effects in hemorrhagic and traumatic shock, respectively.

6. *Glandular secretion*. The ability of antihistamine drugs to antagonize the actions of histamine on the vascular system and on smooth muscle in several locations suggests that such drugs might inhibit lacrimal, salivary, gastric and pancreatic secretion following stimulation of these glands with histamine. If so, they could be employed advantageously to elucidate the question as to whether histamine plays any physiological rôle in these secretory processes. Whether histamine is related to the secretory response of the stomach to food and the secretion occurring in abnormal conditions is a moot question (51, 73, 87). If hypersecretion and gastric retention are etiological or aggravating factors in "peptic ulcer," a drug with gastric secretory depressant and antispasmodic properties would certainly warrant clinical trial.

Loew and Chickering (106) found that subcutaneous injection of 929 F in dogs with gastric pouches (Heidenhain) previous to a subcutaneous injection of histamine actually increased the volume and acidity of secretion over that which obtained during control periods when only the histamine stimulus was employed. Hallenbeck (75) corroborated these findings in experiments in which the histamine stimulus was effected in a more physiological manner, viz.

by slow intravenous infusion or absorption from a beeswax mixture. By an unknown mechanism, 929 F did diminish secretion in response to a meal stimulus in dogs with innervated and denervated pouches (Pavlov, Heidenhain, transplanted).

In dogs with denervated pouches, 1571 F failed to modify secretion after a food stimulus (18) but inhibited secretion in dogs with innervated pouches (75), which suggests that the drug exerted a central nervous action. Ignorance of the fact that secretion may vary according to whether the gastric pouch is innervated or denervated undoubtedly accounts for the statement (146) that, "The effect of compounds 929 F and 1571 F on gastric responses to histamine or on (*sic*) food ingestion in dogs varies with the experimenters." 1571 F almost halved the output of histamine-induced secretion (18) although when large single doses of histamine were employed (27) neither 929 F nor 1571 F inhibited secretion. In summarizing, it is apparent that 929 F failed to antagonize the gastric secretogogue action of histamine and that secretory inhibition demonstrated with 1571 F was probably not referable to a specific antagonism of histamine since pilocarpine-induced secretion from denervated pouches was also inhibited (18).

Pretreatment of dogs with Benadryl inhibited secretion from denervated gastric pouches approximately 40 per cent when the histamine stimulus (0.5 mgm. histamine diphosphate) was given as a single subcutaneous injection (112). However, secretion was not inhibited in all dogs and recent evidence indicates that Benadryl (67, 160) and Pyribenzamine failed to antagonize the secretogogue action of histamine in dogs with various types of pouches. Some inhibition was noted after treatment with several quaternary nitrogen derivatives of Benadryl (69). Benadryl did not prevent formation of duodenal and gastric ulcers in dogs which received large amounts of histamine in beeswax (67). No conspicuous inhibition of histamine-induced gastric secretion occurred in dogs pretreated with diethylaminoethyl-dihydroanthracene (100), diethylaminoethyl fluorene-9-carboxalate (99) or several alkyloxytriazines (105) which weakly antagonized histamine (108). Thiodiphenylamine derivatives (3015 R.P. and 3277 R.P.) apparently failed to inhibit salivary, pancreatic and gastric secretion (80).

In the rat, Neoantergan did not inhibit the secretogogue action of histamine (50 mgm./kgm., subcutaneously), although atropine (200 mgm./kgm., subcutaneously) was effective (24). Seemingly the rat is a poor choice of experimental animal if such large doses of drugs are required to stimulate or depress secretion.

It was stated (164) that the previous ingestion of Antergan by human subjects did not modify secretion of acid following injection of histamine. In six subjects (37) the intravenous injection of Antergan in doses sufficient to produce definite symptoms reduced the volume but not the total amount of hydrochloric acid secreted following subcutaneous injection of histamine dihydrochloride (1.0 mgm.), although the local and systemic effects were attenuated. Neoantergan also failed to reduce the total acid secreted by two subjects, although it eliminated all other effects of histamine. Preliminary studies in man did not provide

unequivocal evidence that Benadryl appreciably antagonized the secretogogue action of histamine (123, 124) and of alcohol (125), which may liberate histamine. Appreciable doses of Benadryl infused intravenously failed to significantly decrease the volume and free acid concentration of gastric juice in duodenal ulcer patients given a test meal, or histamine, subcutaneously or intravenously (133).

The available evidence from man and animals justifies the conclusion that none of the antihistamine drugs directly antagonized the gastric secretogogue action of histamine. Thus, at least one of the prominent effects of histamine was not significantly diminished by antihistamine drugs.

Halpern (77) has stated that the intravenous injection of 2325 R.P. and Antergan in dogs augmented salivary secretion and that thiodiphenylamines (80) did not affect salivary, pancreatic and gastric secretion. In anesthetized cats, Yonkman *et al.* (190) noted that salivary secretion induced by carotid injections of histamine was highly variable but appeared to be diminished or annulled by Pyribenzamine. Lacrimation induced with histamine was stated to be definitely suppressed after injections of Pyribenzamine. Salivary secretion was still elicited by faradization of the chorda tympani nerve or by injecting pilocarpine. In some experiments, Pyribenzamine appeared to sensitize the salivary secretory response to such a variety of agents and procedures as to suggest that experimental conditions were improperly controlled. Among patients receiving therapeutic doses of Antergan, Pyribenzamine or Benadryl there is such a low incidence of xerostomia (163, 146) as to suggest that neither the antihistamine nor atropine-like qualities of these agents are physiologically important in relation to salivary secretion.

Because of the absence or paucity of data and the non-quantitative nature of that which exists, it would be premature to express an opinion concerning the ability of antihistamine drugs to antagonize the secretogogue action of histamine on lacrimal, salivary and pancreatic glands. Since histamine-induced gastric secretion was not antagonized, consideration must be given to the possibility that histamine acts on secretory cells through some mechanism which differs from that obtaining in endothelial and smooth muscle cells; or, antihistamine drugs do not reach or have no affinity for the elements in the secretory cell which react to histamine.

Effect of antihistamine drugs in anaphylaxis and allergy. It is not the prime purpose of this review to present a detailed, critical evaluation of the evidence relating to the histamine theory of anaphylaxis and allergy (cf. 10, 32, 44, 45, 60, 149). The problem is of extreme importance, however, both from theoretical and practical standpoints. Therefore, at least a portion of pertinent evidence now available relative to the use of antihistamine drugs in elucidating the problem is cited.

The symptoms and reactions to histamine in a given species are nearly identical to those occurring in anaphylaxis. The reactions to histamine differ from one species to another since different shock organs or tissues give rise to the dominant reactions. Correspondingly, the anaphylactic reactions differ from

species to species, but in a given species they are always strikingly similar to histamine reactions. Antihistamine drugs should therefore be effective in diminishing histamine and anaphylactic reactions in each species if histamine is significantly involved in anaphylaxis.

All antihistamine drugs which so effectively antagonize the bronchioconstrictive action of histamine are capable of diminishing the severity of anaphylaxis in the guinea pig, in which animal bronchioconstriction is the prominent feature (for refs. cf. 32, 44, 60, 183) of anaphylaxis. Anaphylaxis has been alleviated in intact guinea pigs by treating with 929 F (156, 169, 170), 1571 F (109, 169, 186), Antergan and 2325 R.P. (77, 103), Neoantergan (24), Benadryl (65, 107, 165), Pyribenzamine (7, 65, 118, 122), Antistine (127), Hetramine (59), the thiodiphenylamine, 3015 R.P. (79), and α-naphthylmethylethyl-β-chloroethylamine (1). Tolerated doses of antihistamine drugs are more effective in the prevention of both histamine and anaphylactic shock than comparatively large and even poorly tolerated doses of atropine (77), theophylline (52, 100, 107, 110, 111), papaverine (62, 107, 110, 111, 169) epinephrine (77, 90, 100, 111) and other drugs possessing antispasmodic action (77, 100, 142, 169). All of this evidence, and especially that obtained with antihistamine drugs, supports the belief that histamine is involved in producing the symptoms of anaphylaxis.

With antihistamine drugs, the lowest doses (rarely proven to constitute minimal effective doses) used to demonstrate anti-anaphylactic action in guinea pigs are sometimes greater than the minimal effective doses required to relieve histamine shock. For example, the doses of Pyribenzamine, Neoantergan and especially 3015 R.P. used to demonstrate anti-anaphylactic action (24, 65, 79) were appreciably greater than one would expect necessary in view of their effectiveness in antagonizing administered histamine. Before this type of data can be used as evidence that factors other than histamine are importantly involved in anaphylaxis it will be necessary to obtain quantitative data relating to anaphylactic and histamine shock under conditions as strictly comparable as possible. If the protective action of antihistamine drugs is satisfactorily demonstrated to be appreciably less against anaphylactic shock than against histamine shock it will still not constitute weighty evidence that some factor(s) other than histamine is involved to an important degree since in anaphylaxis histamine may be released in more intimate contact with the effector portion of the reacting cells (100) than can be effected by administering histamine.

Anaphylaxis in the dog is intimately connected with the release and action of histamine (for refs. cf. 32, 44, 60, 183), the most conspicuous feature being hypotension resulting from vasodilatation, increased capillary permeability, contraction of hepatic veins and engorgement of tissues drained by the portal system. In a limited number of experiments, Antergan (77) and Pyribenzamine (191) diminished hypotension and symptoms of anaphylaxis in dogs. Wells, Morris and Dragstedt (183) demonstrated absence of mortality upon injection of horse serum in 22 sensitized dogs treated with Benadryl, whereas 9 or 34.6 per cent of the 26 untreated animals died. The persistence of some degree of hypotension when shock was induced in Benadryl-treated animals does not prove that agents other than histamine were involved since histamine blood

levels were increased and it was previously shown (182) that Benadryl diminished but did not abolish the hypotensive action of the quantities of histamine involved.

In rabbits, anaphylaxis or histamine administration causes transient hypertension and then hypotension and death due to circulatory failure which is secondary to a marked constriction of pulmonary blood vessels (44). Atropine, epinephrine and ephedrine (177) failed to diminish the hypotension occurring during anaphylaxis. Neoantergan (24, 143) proved more effective than 929 F, 1571 F and Antergan in diminishing the hypertensive responses to histamine in non-anesthetized rabbits. In anaphylaxis, 929 F (178) and the sympatholytic agents, 933 F and 883 F, were somewhat effective in diminishing reactions (174, 175) although appreciable vasodilatation and hypotension induced with these agents (41) may account, in part, for the effects. Additional data regarding the effectiveness of potent antihistamine agents, and other drugs, in controlling anaphylaxis in rabbits are needed. Comparisons with the effectiveness of aminophylline and a study of the mechanisms involved would be valuable in view of the possibility that the effectiveness of aminophylline in "epinephrine-fast" asthmatic patients may be due to its ability to diminish pulmonary vascular resistance (100).

Following publication of the statement that Neoantergan actually increased the lethal action of histamine in mice (24 and cf. 80), Mayer and Brousseau (119) compared the effects of other antihistamine drugs in anaphylactic shock and histamine shock. Although Pyribenzamine reduced the severity of anaphylaxis produced with large intravenous doses of antigen, these authors stated that Pyribenzamine and Benadryl were both definite *synergists* of histamine in mice. However, it appears probable that the synergism, or additive effect, was actually between the antihistamine drugs and the large volumes of acidic solution injected intravenously into such small animals (0.25 to 0.75 ml. of 2.0 per cent histamine acid phosphate per mouse; 1.0 ml. per minute). These results cannot be accepted as proof of an exception to the parallelism between histamine and anaphylactic shock in a given animal species.

Further evidence concerning the effectiveness of antihistamine drugs in diminishing or preventing anaphylactic responses is included in the section dealing with localized cutaneous reactions to histamine and antigens (vide supra) in animals and man.

A few studies relative to anaphylaxis *in vitro* have been made. The spasm of sensitized smooth muscle which follows addition of antigen to the perfusing or bath fluid is likely due to histamine liberated from the tissue. For the most part the reactions are qualitatively and quantitatively typical of those induced by adding proper amounts of histamine.

Bronchospasm induced by adding antigen to fluid perfused through isolated, sensitized, guinea-pig lungs was lessened by Benadryl (49), whereas Pyribenzamine (192) failed to decrease spasm, possibly because large amounts of antigen (10 mgm. of horse serum) were employed. Perfusion experiments with isolated lungs are not ideally suited for quantitative determinations of the relative potencies of antihistamine drugs or for determination of quantitative relation-

ships referable to degree of effect against antigen and histamine. 1571 F (186) prevented antigen and histamine from decreasing the flow through coronary vessels of the sensitized guinea pig-heart.

The spasm induced by adding antigen to isolated intestinal and uterine muscle from sensitized animals has been diminished or annulled by 929 F (156, 169, 170), Antergan and 2325 R.P. (77), 3015 R.P. (79) and Antistine (127).

Reactions to trypsin. In view of the similarity of reactions in trypsin shock and anaphylaxis, as well as the numerous observations which supported the hypothesis that the toxicity of trypsin was referable to the enzymatic release of histamine from tissues, it is significant that Benadryl 1, failed to provide protection against the lethal action of trypsin in guinea pigs and dogs; 2, only very slightly diminished the hypotensive action of trypsin in dogs (184), and 3, did not prevent the increased capillary permeability induced with crystalline trypsin in rabbit's skin (98). Furthermore, trypsin injections in dogs did not increase blood histamine sufficiently to account for severity of reactions (184), and in view of other evidence (171) it is now doubtful that the histamine hypothesis, relative to the toxicity of trypsin, is tenable.

Summary. Thus the evidence cited indicates that specific antihistamine drugs diminish several pharmacological responses to administered histamine which are of the type which could account for a variety of anaphylactic reactions. These drugs also diminish those anaphylactic responses which are considered as being due to histamine. Furthermore, in several animal species a parallelism exists between histamine and anaphylactic responses, both of which are prevented or controlled with antihistamine drugs. The use of specific antihistamine drugs as pharmacological tools has therefore provided evidence which even more firmly establishes the concept that histamine causes the major manifestations of anaphylaxis.

In addition, the effectiveness of antihistamine drugs in controlling certain allergic manifestations in man has been established and it is pertinent to note that therapeutic efficacy demonstrated thus far is restricted to allergic diseases, and possibly a few other conditions in which there are some reasons for considering histamine as an etiological agent. This constitutes strong evidence that histamine is involved to an appreciable extent in several allergic diseases.

Antihistamine drugs have been selected and partially evaluated on the basis of their striking effectiveness in alleviating bronchioconstriction induced in guinea pigs with histamine or by anaphylaxis. In clinical use these agents have exhibited much less efficacy in bronchial asthma than in angioneurotic edema, urticaria and seasonal rhinitis, i.e., in conditions in which vascular reactions undoubtedly play a prominent rôle. Limitation of space does not permit discussion of possible explanations of such differences, although mention should be made of the possibility that bronchiolar smooth muscle of the guinea pig differs markedly from that in other species, both in regard to sensitivity to histamine and the blocking action of antihistamine drugs.

Since antihistamine drugs do not always completely block effects of histamine and fail to diminish significantly certain effects such as the stimulant action of

histamine on gastric secretory cells, their failure to diminish or prevent a given response does not necessarily eliminate the possibility that histamine bears a causal relationship.

SPECIFICITY OF ANTIHISTAMINE DRUGS AND RELATION TO AUTONOMIC FUNCTIONS. From the experimental evidence available it appears that 1571 F, Antergan, Neoantergan, Pyribenzamine and Antistine exert a low, and probably insignificant, degree of atropine-like action. It is doubtful whether the antihistamine and anti-allergic activity of Benadryl is closely related to atropine-like action since such action is weaker (112, 187) than that exerted by antispasmodic drugs such as Pavatrine and Trasentin which do not prevent anaphylaxis or effects of histamine on bronchioles and the vascular system. In support of the belief that atropine-like action of antihistamine drugs is unimportant is the fact that those drugs employed clinically seldom produce tachycardia or inhibit salivation. From a theoretical standpoint, even a slight degree of atropine-like activity could be important if the drugs simultaneously enhanced activity of tissues receiving sympathetic innervation. Consideration must therefore be given to evidence that antihistamine drugs potentiate responses to epinephrine and stimulation of adrenergic nerves.

Antergan enhanced retraction of the nictitating membrane in rabbits and the rise of blood pressure in dogs following injections of epinephrine (40, 137). When adequate doses were injected intravenously, Benadryl, Neoantergan and Pyribenzamine increased the height and duration of the pressor response to epinephrine in anesthetized dogs (112, 167, 168, 189). Sherrod *et al.* (167, 168) demonstrated that this phenomenon was not referable to an atropine-like blocking action of the antihistamine drugs on the cardiac vagus since under the experimental conditions even atropine failed to enhance the pressor action of epinephrine, and furthermore, the antihistamine drugs still induced the phenomenon in atropinized and vagotomized dogs. The biological significance of this potentiation of epinephrine must be questioned until the phenomenon is demonstrated in unanesthetized animals or man and data are adduced to show that other adrenergic responses are appreciably augmented by antihistamine drugs, and that this property is significantly related to antihistamine action. The failure of Benadryl to elevate the blood sugar level in man or to decrease glucose tolerance (126, 144) strongly suggests that no appreciable sympathomimetic action was exerted since a sensitive indicator of epinephrine action is a rise in blood sugar (11). In cats, Pyribenzamine (189, 190) did not consistently enhance several responses to injected epinephrine and sympathetic nerve stimulation. No potentiation of adrenergic responses has been reported to occur after administration of 929 F, 1571 F, and 2325 R.P.

There is still other evidence that adrenergic potentiation or sympathomimetic action does not account for the pronounced antagonism of histamine effected by the drugs under consideration. There is agreement (77, 86, 111) that ephedrine fails to reduce the severity of bronchioconstriction induced with large doses of histamine in guinea pigs, although it did relieve bronchioconstriction in an asthmatic patient injected with histamine (35). Doses of epinephrine which

produce symptoms in guinea pigs are not as effective as tolerated doses of several antihistamine drugs in alleviating symptoms of severe histamine shock (77, 80, 111). The author was unable to demonstrate any synergism between ephedrine and Benadryl in reducing mortality of guinea pigs upon induction of histamine shock (105). A strong argument against the hypothesis that sympathomimetic properties, or some type of adrenergic potentiation, is a requirement for potent antihistamine action resides in the fact that 929 F antagonizes histamine as well as certain effects of epinephrine (22, 31); also there is the recent demonstration (1) that certain adrenergic blocking agents are also exceedingly potent antagonists of histamine, both actions being rapid in onset and of remarkable duration.

α-Naphthylmethylethyl-β-chloroethylamine diminished or blocked adrenergic excitatory actions induced by nerve stimulation or epinephrine injection in dogs and cats. Pressor responses to epinephrine were readily reversed. On a dosage basis, this naphthyl derivative was more effective than epinephrine in preventing histamine shock and *anaphylactic shock* in guinea pigs. In dogs, the hypotensive action of histamine was blocked or diminished. The antihistamine action could not be entirely dependent on sympatholytic or adrenergic blocking action since among higher alkyl homologues of the naphthyl compound there were individual compounds with singular sympatholytic potency which did not antagonize histamine. Furthermore, the author (105) has been unable to demonstrate antagonism of histamine with adrenergic blocking agents such as Priscol (83), dibenzyl-β-chloroethylamine (Dibenamine) (134) and a series of benzhydrylalkyl-β-chloroethylamines (113). The dioxane derivatives, 2-piperidinomethyl-1, 4-benzodioxane (933 F) and 2-diethylaminomethyl-1, 4-benzodioxane (883 F), which are best known as adrenergic blocking drugs, were somewhat effective in diminishing effects of injected histamine, as well as anaphylaxis, in rabbits (174, 175). Such effects could be related to vasodilatation and hypotension produced by these compounds (41). 933 F and 883 F failed to decrease severity of histamine shock in guinea pigs (105), although 933 F diminished histamine-induced intestinal spasm (173). If some potent antihistamine drugs actually do potentiate adrenergic responses whereas others exert a sympatholytic action, it is difficult to conceive that strong antihistamine action is dependent on either type of activity.

From the evidence now available it appears unlikely that the anti-anaphylactic and anti-allergic activity of Antergan, Neoantergan, Benadryl, Pyribenzamine and Antistine is importantly related to either an atropine-like activity or to an augmentation of adrenergic responses. These drugs can be regarded as fairly specific antagonists of histamine. Further studies relating to specificity of action are needed and the detailed study of all antihistamine compounds should obviously include determination of atropine-like, sympathomimetic, sympatholytic, quinidine-like, local anesthetic, antispasmodic and other properties.

Mode of Action. Mode of drug antagonism may be of three types (30, 68, 148). An example of indirect physiological antagonism is the ability of epinephrine to induce biological responses which are independent of, and dia-

metrically opposed to, such actions of histamine as the production of bronchospasm, vasodilatation, increased capillary permeability, contraction of intestinal muscle, etc. This type of antagonism is rather non-specific in character since epinephrine will diminish or annul certain pharmacological effects of numerous drugs. The other two types are both direct antagonisms. Drug action may be annulled or diminished if direct chemical reactions render the drug inactive by reduction, oxidation, conjugation or formation of salts or complexes. Finally, a direct antagonism may involve competition between two substances for a given site of action or receptive substance. This type of antagonism is reversible and one of the most specific in nature.

It has been repeatedly stated that there is no indication that antihistamine drugs act by combining with or destroying histamine, or by activating diamine oxidase or "histaminase" which are known to inactivate histamine and other amines, although the only pertinent evidence reported concerns 929 F (31, 169).

The major action of antihistamine drugs is probably exerted directly on the peripheral effector cells which respond to histamine since antagonism has frequently been demonstrated on isolated tissues or organs. The antagonism exerted peripherally could be modified by nervous or humoral influences in the intact animal. Quantitative data obtained by Wells *et al.* (181, 182) indicates the probability that Benadryl antagonizes the hypotensive action of histamine in dogs by adsorbing onto the site of action of histamine where it exerts no action except that of competing with and diminishing the action of histamine. The quantitative data reported by Halpern (81) concerning Antergan-histamine antagonism on isolated intestinal tissue also renders it probable that a direct, competitive inhibition was involved. The available evidence suggests that the antagonism of histamine is comparable to the antagonism between atropine and acetylcholine, and between adrenergic blocking agents and epinephrine, where there is a competition between two types of molecules (30, 148, 182).

Although antagonism of histamine with drugs may be regarded as a competitive affair it does not necessarily involve compounds with close chemical structure as is the case with chemical analogues in the p-aminobenzoic acid-sulfonamide and other metabolite-antimetabolite systems (cf. 180). Quaternary ammonium derivatives of Benadryl are choline ethers which possess, in addition to antihistamine action, the unique property of antagonizing acetylcholine (109, 112, 187). It is doubtful whether this choline ether-acetylcholine antagonism is closely related to analogous chemical structures since the most pronounced action was the choline ether-histamine antagonism where no similarity in structure obtains. Since Bovet and Walthert (24) first noted the diversity of the chemical nature of antihistamine drugs and the homogeneity of their pharmacological actions, several other types of antihistamine drugs have been found (1, 80, 110). The diverse chemical nature of compounds which antagonize histamine, along with the evidence that some of these compounds also alter responses to acetylcholine and epinephrine, immediately suggests that their antihistamine action is not closely related to a competition between chemical analogues. It would appear much more likely that antihistamine drugs alter surface tension or

are adsorbed onto one or more components of the reactive tissue so as to obviate the usual response to histamine. The specificity of antihistamine action varies with different compounds for some alter the response to acetylcholine, epinephrine and possibly other substances.

Rocha e Silva (151) pointed out the fact that there is a possible interdependence between the sites of attack of acetylcholine and histamine, since adequate quantities of atropine, the specific inhibitor of acetylcholine, also antagonize some of the effects of histamine (cf. 33). Other facts support the concept that acetylcholine and histamine act on closely related cellular components or on identical components in a slightly different manner. First, there is the close similarity of action of histamine and acetylcholine such as their secretory and vasodilator effects and the spasmogenic action on uterine, bronchiolar and intestinal smooth muscle. Secondly, some anti-acetylcholine activity is exhibited by most of the agents which antagonize actions of histamine, including the antihistamine drugs herein discussed, amino-acids (cf. 78, 142, 150, 151), certain derivatives of histamine and amino-acids (20, 150, 151), and a number of antispasmodics (52, 100, 108, 112, 161).

Consideration should probably also be given to an interdependence between the sites of attack of histamine and epinephrine for there are some similarities in their action, as well as conspicuous differences. It should be noted, however, that most of the antihistamine drugs, under certain conditions and with proper doses, either enhance or diminish certain effects of epinephrine. The three agents, acetylcholine, epinephrine, and histamine, all occur in the animal organism and may exert their physiological activity by acting through closely related mechanisms.

Some effects of epinephrine are diminished or blocked by 929 F (22, 31). Dioxane derivatives (883 F and 933 F) which block or reverse certain actions of epinephrine (41, and cf. 72) may antagonize histamine (169, 173–175), although the evidence is not entirely convincing. Of especial significance is the recent demonstration that α-naphthylmethylethyl-β-chloroethylamine, and related compounds, strongly antagonize certain actions of both histamine and epinephrine (1). Elucidation of the pharmacological properties and mode of action of compounds which antagonize both histamine and epinephrine in an extraordinary manner will indeed be important since comparisons can be made with results obtained with those antihistamine compounds (Antergan, Neoantergan, Benadryl and Pyribenzamine) which, under experimental conditions at least, are known to enhance the pressor and possibly other actions of epinephrine.

The effectiveness of antihistamine drugs in controlling anaphylaxis and allergy appears to be definitely related to their ability to diminish or block the effects of histamine upon vascular and visceral smooth muscle and permeability of capillaries. None of the evidence suggests that antihistamine drugs influence the antigen-antibody reaction. It is pertinent to note that blood levels of histamine increased in the usual manner following injection of antigen in dogs protected with Benadryl (183). The rise in blood histamine reported as occurring in normal human subjects following treatment with Antergan and Neoanter-

gan (140) requires confirmation and explanation. In animals, Antistine (128) and Antergan (104) did not inhibit antibody formation. The injection of Pyribenzamine (7) failed to change the complement titer of serum, and when the drug was mixed with antibody and antigen the precipitin titer was not altered.

REFERENCES

(1) Achenbach, P. and E. R. Loew. Fed. Proc. 6: 304, 1947.
(2) Ackerman, D. and H. Mauer. Pflüger's Arch. 247: 623, 1944.
(3) Ahlmark, A. Acta physiol. Scandinav. 9: 5, suppl't. 28, 1944.
(4) Alexander, F. Quart. J. Exper. Physiol. 33: 71, 1944.
(5) Anrep, G. V., G. S. Barsoum, S. Salama and Z. Souidan. J. Physiol. 103: 297, 1944.
(6) Arbesman, C. E., G. F. Koepf and A. R. Lenzner. J. Allergy 17: 275, 1946.
(7) Arbesman, C. E., G. F. Koepf and G. E. Miller. J. Allergy 17: 203, 1946.
(8) Baer, R. L. and M. B. Sulzberger. J. Invest. Dermat. 7: 201, 1946.
(9) Barnett, S. E., F. M. Barbas and S. B. Goss. J. Mich. State Med. Soc. 45: 771, 1946.
(10) Best, C. H. and E. W. McHenry. Physiol. Rev. 11: 371, 1931.
(11) Beyer, K. H. Physiol. Rev. 26: 169, 1946.
(12) Bier, O. and M. Rocha e Silva. Compt. rend. Soc. de biol. 129: 769, 1938.
(13) Blum, H. F. Physiol. Rev. 25: 483, 1945.
(14) Bonvallet, M. and P. Decourt. Compt. rend. Soc. de biol. 138: 224, 1944.
(15) Bonvallet, M. and P. Decourt. Compt. rend. Soc. de biol. 138: 305, 1944.
(16) Boquet, A. Ann. Inst. Pasteur 69: 55, 1943.
(17) Boquet, P. Compt. rend. Soc. de biol. 137: 347, 1943.
(18) Bourque, J. E. and E. R. Loew. Am. J. Physiol. 138: 341, 1943.
(19) Bourquin, J. B. Schweiz. med. Wchnschr. 76: 296, 1946.
(20) Bovet, D., R. Horclois and J. Fournel. Compt. rend. Soc. de biol. 138: 165, 1944.
(21) Bovet, D., R. Horclois and F. Walthert. Compt. rend. Soc. de biol. 138: 99, 1944.
(22) Bovet, D. and P. Maderni. Compt. rend. Soc. de biol. 114: 980, 1933.
(23) Bovet, D. and A.-M. Staub. Compt. rend. Soc. de biol. 124: 547, 1937.
(24) Bovet, D. and F. Walthert. Ann. pharmaceutiques françaises 2: suppl. to no. 4, 1, 1944.
(25) Brack, W. Schweiz. med. Wchnschr. 76: 316, 1946.
(26) Breton, A. Compt. rend. Soc. de biol. 137: 254, 1943.
(27) Burchell, H. B. and R. L. Varco. J. Pharmacol. and Exper. Therap. 75: 1, 1942.
(28) Celice, J., M. Perrault and P. Durel. Bull. et mém. Soc. méd. d. hôp. de Paris 58: 284, 1942.
(29) Chauchard, B. and P. Chauchard. Compt. rend. Soc. de biol. 137: 708, 1943.
(30) Clark, A. J. The mode of action of drugs on cells. Williams & Wilkins Company, Baltimore, 1933.
(31) Climenko, D. R., R. Homburger and F. H. Messer. J. Lab. and Clin. Med. 27: 289, 1941.
(32) Code, C. F. Ann. Allergy 2: 457, 1944.
(33) Code, C. F. Proc. Staff Meet. Mayo Clin. 20: 439, 1945.
(34) Crammer, J. L. and M. P. Helle. Nature 154: 18, 1944.
(35) Curry, J. J. J. Clin. Investigation 25: 792, 1946.
(36) Danielopolu, D., M. Popesco and E. Mezincesco. Compt. rend. Soc. de biol. 138: 380, 1944.
(37) Decourt, J., A. Rinieri and P. Sonnet. Compt. rend. Soc. de biol. 139: 470, 1945.
(38) Decourt, M. J. Semaine d. hôp. Paris 21: 707, 1945.
(39) Decourt, P. and M. Mollaret. Bull. et mém. Soc. méd. d. hôp. de Paris 58: 265, 1942.
(40) DeCuyper, T. Arch. internat. de pharmacodyn. et de thérap. 72: 360, 1946.

(41) DeVleeschhouwer, G. Arch. internat. de pharmacodyn. et de thérap. 50: 251, 1935.
(42) Dews, P. B. and J. D. P. Graham. Brit. J. Pharmacol. 1: 278, 1946.
(43) Dragstedt, C. A. Quart. Bull., Northwestern Univ. M. School 14: 283, 1940.
(44) Dragstedt, C. A. Physiol. Rev. 21: 563, 1941.
(45) Dragstedt, C. A. Quart. Bull., Northwestern Univ. M. School 17: 102, 1943.
(46) Dragstedt, C. A. Quart. Bull., Northwestern Univ. M. School 19: 303, 1945.
(47) Dragstedt, C. A. and F. B. Mead. Proc. Soc. Exper. Biol. and Med. 32: 1435, 1935.
(48) Elias, H. and E. H. McGavack. Proc. Soc. Exper. Biol. and Med. 61: 133, 1946
(49) Ellis, F. W. Fed. Proc. 4: 117, 1945.
(50) Ellis, F. W. and J. F. Newsome. Fed. Proc. 5: 176, 1946.
(51) Emmelin, N. and G. S. Kahlson. Acta physiol. Scandinav. 8: 289, 1944.
(52) Emmelin, N., G. S. Kahlson and K. Lindstrom. Acta physiol. Scandinav. 3: 39, 1941.
(53) Emmelin, N., G. S. Kahlson and F. Wicksell. Acta physiol. Scandinav. 2: 123, 1941.
(54) Essex, H. E. Physiol. Rev. 25: 148, 1945.
(55) Eyermann, C. H. J. Allergy 17: 210, 1946.
(56) Feinberg, S. M. J. Allergy 17: 217, 1946.
(57) Feinberg, S. M. and S. Friedlaender. J. Allergy 16: 296, 1945.
(58) Feinberg, S. M. and S. Friedlaender. Am. J. Med. Sci. 213: 58, 1947.
(59) Feinstone, W. H., R. D. Williams and B. Rubin. Proc. Soc. Exper. Biol. and Med. 63: 158, 1946.
(60) Feldberg, W. Ann. Rev. Physiol. 3: 671, 1941.
(61) Fischer, E. Physiol. Rev. 24: 467, 1944.
(62) Frank, D. E. J. Immunol. 52: 59, 1946.
(63) Friedlaender, A. S. and S. Friedlaender. J. Lab. and Clin. Med. 31: 1350, 1946.
(64) Friedlaender, S. and S. M. Feinberg. J. Allergy 17: 129, 1946.
(65) Friedlaender, S., S. M. Feinberg and A. R. Feinberg. Proc. Soc. Exper. Biol. and Med. 62: 65, 1946.
(66) Friedlaender, S., S. M. Feinberg and A. R. Feinberg. J. Lab. and Clin. Med. 32: 47, 1947.
(67) Friesen, S. R., I. D. Baronofsky and O. H. Wangensteen. Proc. Soc. Exper. Biol. and Med. 63: 23, 1946.
(68) Gaddum, J. H. Trans. Faraday Soc. 39: 323, 1943.
(69) Garrido, G. K. The effect of antihistamine, atropine-like and sympathomimetic compounds upon histamine-induced gastric secretion. Thesis, Northwestern Univ. Med. School, 1946.
(70) Gelvin, E. P., H. Elias and T. H. McGavack. J. Pharmacol. and Exper. Therap. 89: 101, 1947.
(71) Gelvin, E. P. and T. H. McGavack. Bull. N. Y. Med. Col., Flower and Fifth Ave. Hospitals 9: 51, 1946.
(72) Goodman, L. and A. Gilman. The pharmacological basis of therapeutics. The Macmillan Company, New York, 1941.
(73) Grossman, M. I. and A. C. Ivy. Gastroenterology 7: 134, 1946.
(74) Guggenheim, M. Die biogenen amine. Karger, Basle, 1940.
(75) Hallenbeck, G. A. Am. J. Physiol. 139: 329, 1943.
(76) Hallenbeck, G. A., C. F. Code and F. C. Mann. Gastroenterology 1: 588, 1943.
(77) Halpern, B. N. Arch. internat. de pharmacodyn. et de thérap. 68: 339, 1942.
(78) Halpern, B. N. Compt. rend. Soc. de biol. 139: 625, 1945.
(79) Halpern, B. N. Compt. rend. Soc. de biol. 140: 363, 1946.
(80) Halpern, B. N. and R. Ducrot. Compt. rend. Soc. de biol. 140: 361, 1946.
(81) Halpern, B. N. and G. Mauric. Compt. rend. Soc. de biol. 140: 440, 1946.
(82) Harris, R., T. H. McGavack and H. Elias. J. Lab. and Clin. Med. 31: 1148, 1946.

(83) Hartmann, M. and H. Isler. Arch. f. exper. Path. u. Pharmakol. 192: 141, 1939.
(84) Hewitt, W. L. and J. J. Curry. J. Lab. and Clin. Med. 32: 42, 1947.
(85) Ingraham, R. C. and H. C. Wiggers. Fed. Proc. 5: 50, 1946.
(86) Issekutz, B. V. and P. Genersich. Arch. f. exper. Path. u. Pharmakol. 202: 201, 1943.
(87) Ivy, A. C. Trans. and Stud., Coll. Physicians, Philadelphia 12: 101, 1944.
(88) Jourdan, F. and J. Chatonnet. Compt. rend. Soc. de biol. 137: 559, 1943.
(89) Kallós, P. and L. Kallós-Deffner. Acta med. Scandinav. 116: 409, 1944.
(90) Kallós, P. and L. Kallós-Deffner. Acta med. Scandinav. 116: 440, 1944.
(91) Kallós, P. and W. Pagel. Acta med. Scandinav. 91: 292, 1937.
(92) Kapeller-Adler, R. Biochem. J. 35: 213, 1941.
(93) Kapeller-Adler, R. Biochem. J. 38: 270, 1944.
(94) Kellaway, C. H. Ann. Rev. Biochem. 8: 541, 1939.
(95) Koepf, G. F., C. E. Arbesman and C. Munafo. J. Allergy 17: 271, 1946.
(96) Kwiatkowski, H. J. Physiol. 102: 32, 1943.
(97) Landis, E. M. Physiol. Rev. 14: 404, 1934.
(98) Last, M. R. and E. R. Loew. J. Pharmacol. and Exper. Therap. 89: 81, 1947.
(99) Lehmann, G. and P. K. Knoefel. J. Pharmacol. and Exper. Therap. 74: 217, 1942.
(100) Lehmann, G. and J. W. Young. J. Pharmacol. and Exper. Therap. 83: 90, 1945.
(101) Levin, S. J. J. Allergy 17: 145, 1946.
(102) Lewis, T. The blood vessels of the human and their responses. Shaw and Sons, London, 1927.
(103) Leya, A. Compt. rend. Soc. de biol. 140: 191, 1946.
(104) Leya, A. Compt. rend. Soc. de biol. 140: 194, 1946.
(105) Loew, E. R. Unpublished data.
(106) Loew, E. R. and O. Chickering. Proc. Soc. Exper. Biol. and Med. 48: 65, 1941.
(107) Loew, E. R. and M. E. Kaiser. Proc. Soc. Exper. Biol. and Med. 58: 235, 1945.
(108) Loew, E. R., M. E. Kaiser and M. Anderson. J. Pharmacol. and Exper. Therap. 86: 7, 1946.
(109) Loew, E. R., M. E. Kaiser and M. Anderson. Fed. Proc. 5: 190, 1946.
(110) Loew, E. R., M. E. Kaiser and V. Moore. J. Pharmacol. and Exper. Therap. 83: 120, 1945.
(111) Loew, E. R., M. E. Kaiser and V. Moore. J. Pharmacol. and Exper. Therap. 86: 1, 1946.
(112) Loew, E. R., R. MacMillan and M. E. Kaiser. J. Pharmacol. and Exper. Therap. 86: 229, 1946.
(113) Loew, E. R., A. Micetich and P. Achenbach. Fed. Proc., 6: 351, 1947.
(114) Lofstrom, J. E. and C. E. Nurnberger. Am. J. Roent. 56: 211, 1946.
(115) MacIntosh, F. C. Quart. J. Exper. Physiol. 28: 87, 1938.
(116) Marshall, P. B. J. Physiol. 102: 180, 1943.
(117) Mathieson, D., H. W. Hays, D. Chess, A. Cameron and F. F. Yonkman. Fed. Proc. 5: 192, 1946.
(118) Mayer, R. L. J. Allergy 17: 153, 1946.
(119) Mayer, R. L. and D. Brousseau. Proc. Soc. Exper. Biol. and Med. 63: 187, 1946.
(120) Mayer, R. L., H. W. Hays, D. Brousseau, D. Mathieson, B. Rennick and F. F. Yonkman. J. Lab. and Clin. Med. 31: 749, 1946.
(121) Mayer, R. L., C. P. Huttrer and C. R. Scholz. Fed. Proc. 4: 129, 1945.
(122) Mayer, R. L., C. P. Huttrer and C. R. Scholz. Science 102: 93, 1945.
(123) McElin, T. W. and B. T. Horton. Proc. Staff Meet., Mayo Clin. 20: 417, 1945.
(124) McElin, T. W. and B. T. Horton. Gastroenterology 7: 100, 1946.
(125) McGavack, T. H., H. Elias and L. J. Boyd. Gastroenterology 6: 439, 1946.
(126) McGavack, T. H., H. Elias and L. J. Boyd. J. Lab. and Clin. Med. 31: 560, 1946.
(127) Meier, R. and K. Bucher. Schweiz. med. Wchnschr. 76: 294, 1946.
(128) Meier, R. and K. Bucher. Experientia 2: 140, 1946.

(129) Menkin, V. J. Exper. Med. 64: 485, 1936.
(130) Menkin, V. J. Exper. Med. 67: 145, 1938.
(131) Menkin, V. Physiol. Rev. 18: 366, 1938.
(132) Minard, D. and S. R. Rosenthal. Proc. Soc. Exper. Biol. and Med. 44: 237, 1940.
(133) Moersch, R. U., A. B. Rivers and C. G. Morlock. Gastroenterology 7: 91, 1946.
(134) Nickerson, M. and L. S. Goodman. J. Pharmacol. and Exper. Therap. 89: 167, 1947.
(135) O'Leary, P. A. and E. M. Farber. Proc. Staff Meet., Mayo Clin. 21: 295, 1946.
(136) Parrot, J. L. Compt. rend. Soc. de biol. 136: 715, 1942.
(137) Parrot, J. L. Compt. rend. Soc. de biol. 137: 378, 1943.
(138) Parrot, J. L. and J. Lefebvre. Compt. rend. Soc. de biol. 137: 316, 1943.
(139) Parrot, J. L. and J. Lefebvre. Compt. rend. Soc. de biol. 137: 756, 1943.
(140) Pellerat, J. and M. Murat. Compt. rend. Soc. de biol. 140: 297, 1946.
(141) Pfeiffer, C. C. Fed. Proc. 5: 197, 1946.
(142) Pfeiffer, C. C. and E. R. Loew. Ann. Rev. Physiol. 9: 651, 1947.
(143) Ramanamanjary, W. Compt. rend. Soc. de biol. 138: 480, 1944.
(144) Reinstein, H. and T. H. McGavack. J. Clin. Endocrinol. 6: 643, 1946.
(145) Rennick, B., D. Chess, H. W. Hays, D. Mathieson, R. L. Mayer and F. F. Yonkman. Fed. Proc. 4: 133, 1945.
(146) Report of Council on Pharmacy and Chemistry. J. A. M. A. 132: 702, 1946.
(147) Rieveschl, G. and O. M. Gruhzit. Fed. Proc. 4: 150, 1945.
(148) Roblin, R. O. Chem. Rev. 38: 255, 1946.
(149) Rocha e Silva, M. Arch. Path. 33: 387, 1942.
(150) Rocha e Silva, M. J. Allergy 15: 399, 1944.
(151) Rocha e Silva, M. J. Pharmacol. and Exper. Therap. 80: 399, 1944.
(152) Rocha e Silva, M. and O. Bier. Compt. rend. Soc. de biol. 129: 773, 1938.
(153) Rocha e Silva, M. and C. A. Dragstedt. Proc. Soc. Exper. Biol. and Med. 46: 303, 1941.
(154) Rocha e Silva, M. and C. A. Dragstedt. J. Pharmacol. and Exper. Therap. 73: 405, 1941.
(155) Rose, B. and J. S. L. Browne. Am. J. Physiol. 124: 412, 1938.
(156) Rosenthal, S. R. and M. L. Brown. J. Immunol. 38: 259, 1940.
(157) Rosenthal, S. R. and D. Minard. J. Exper. Med. 70: 415, 1939.
(158) Rosenthal, S. R., D. Minard and E. Lambert. Proc. Soc. Exper. Biol. and Med. 52: 317, 1943.
(159) Rothman, S. Physiol. Rev. 21: 357, 1941.
(160) Sangster, W., M. I. Grossman and A. C. Ivy. Gastroenterology 6: 436, 1946.
(161) Schaumann, O. Arch. f. exper. Path. u. Pharmakol. 196: 109, 1940.
(162) Schindler, O. Schweiz. med. Wchnschr. 76: 300, 1946.
(163) Schnitzer, A. Dermatologica 91: 92, 1945.
(164) Sciclounoff, F. and R. Junet. Rev. méd. de la Suisse Rom. 63: 570, 1943.
(165) Selle, W. A. Fed. Proc. 5: 93, 1946.
(166) Selle, W. A. Texas Rep. Biol. and Med. 4: 138, 1946.
(167) Sherrod, T., E. R. Loew and H. Schloemer. J. Pharmacol. and Exper. Therap. 89: 247, 1947.
(168) Sherrod, T., H. Schloemer and E. R. Loew. Fed. Proc. 5: 202, 1946.
(169) Staub, A. M. Ann. Inst. Pasteur 63: 400, 485, 1939.
(170) Staub, A. M. and D. Bovet. Compt. rend. Soc. de biol. 125: 818, 1937.
(171) Tagnon, H. J. J. Clin. Investigation 24: 1, 1945.
(172) Todd, L. C. Ann. Allergy 4: 282, 1946.
(173) Ungar, G., J. L. Parrot and D. Bovet. Compt. rend. Soc. de biol. 124: 445, 1937.
(174) Vallery-Radot, P., D. Bovet, G. Mauric and A. Holtzer. Compt. rend. Soc. de biol. 136: 354, 1942.
(175) Vallery-Radot, P., D. Bovet, G. Mauric and A. Holtzer. Compt. rend. Soc. de biol. 136: 335, 1942.

(176) VALLERY-RADOT, P., G. MAURIC AND B. N. HALPERN. Compt. rend. Soc. de biol. **140**: 480, 1946.
(177) VALLERY-RADOT, P., G. MAURIC AND MME. HOLTZER. Compt. rend. Soc. de biol. **137**: 164, 1943.
(178) VALLERY-RADOT, P., G. MAURIC AND A. HOLTZER. Compt. rend. Soc. de biol. **137**: 295, 1943.
(179) WAREMBOURG, H., LINQUETTE AND MACHON. Compt. rend. Soc. de biol. **138**: 417, 1944.
(180) WELCH, A. D. Physiol. Rev. **25**: 687, 1945.
(181) WELLS, J. A. AND H. C. MORRIS. Fed. Proc. **4**: 140, 1945.
(182) WELLS, J. A., H. C. MORRIS, H. B. BULL AND C. A. DRAGSTEDT. J. Pharmacol. and Exper. Therap. **85**: 122, 1945.
(183) WELLS, J. A., H. C. MORRIS AND C. A. DRAGSTEDT. Proc. Soc. Exper. Biol. and Med. **61**: 104, 1946.
(184) WELLS, J. A., H. C. MORRIS AND C. A. DRAGSTEDT. Proc. Soc. Exper. Biol. and Med. **62**: 209, 1946.
(185) WIGGERS, C. J. Physiol. Rev. **22**: 74, 1942.
(186) WILCOX, H. B. AND B. C. SEEGAL. J. Immunol. **44**: 219, 1942.
(187) WINDER, C. V., M. E. KAISER, M. M. ANDERSON AND E. M. GLASSCO. J. Pharmacol. and Exper Therap. **87**: 121, 1946.
(188) WOLFF, H. G. Research Nerv. and Ment. Dis. **23**: 173, 1943.
(189) YONKMAN, F. F., D. CHESS, H. W. HAYS, B. RENNICK AND R. MAYER. Fed. Proc. **5**: 216, 1946.
(190) YONKMAN, F. F., D. CHESS, D. MATHIESON AND N. HANSEN. J. Pharmacol. and Exper. Therap. **87**: 256, 1946.
(191) YONKMAN, F. F., H. W. HAYS AND B. RENNICK. Fed. Proc. **4**: 144, 1945.
(192) YONKMAN, F. F., E. OPPENHEIMER, B. RENNICK AND ELIZABETH PELLET. J. Pharmacol. and Exper. Therap. **89**: 31, 1947.

THE METABOLISM OF ACETIC ACID IN ANIMAL TISSUES

KONRAD BLOCH

Department of Biochemistry and Institute of Radiobiology and Biophysics, University of Chicago

Although acetic acid has long been suspected to be an intermediate in animal metabolism, the significance of its rôle in biochemical processes has only recently been recognized. The lack of interest in acetic acid metabolism during the past is in part attributable to the inadequacy of methods for identification and for quantitative determination of small amounts of acetic acid, and partly due to the fact that in isolated tissues acetic acid appears to possess only limited reactivity.

Until recently the principal approach to problems of intermediary metabolism has been the balance experiment, i.e., the measurement of the changes in concentration of the reaction product induced by varying concentrations of suspected precursor. The only reactions of acetic acid in animal tissues which could be reasonably well established by balance experiments were the acetylation of foreign amines (1, 2) and the formation of acetoacetic acid (3, 4, 5). A much greater variety of reactions involving acetic acid has been revealed with the aid of isotopically labeled substrates. The utility of the tracer technique for the study of intermediary metabolism derives from the fact that with its aid biochemical conversions become demonstrable irrespective of whether or not the total quantity of the reaction product undergoes a change. Thus the use of labeled acetic acid has been of particular service in detecting acetate formation and in demonstrating its participation in processes which proceed independently of exogenous acetate supply, viz., the biological syntheses of steroids, fatty acids, porphyrin, glucose and uric acid.

In general, it will prove more useful to review a broad segment of intermediary metabolism rather than the metabolism of a single compound. Indeed, the great diversity of biochemical processes, catabolic and anabolic, in which acetic acid is involved would require a much more comprehensive discussion of related subjects of intermediary metabolism than can be given here. Also, the present review is concerned mainly with the metabolism of acetic acid in animal tissues, although reference will be made frequently to processes occurring in other cells. This will be necessary because in many cases the discovery of a biochemical process in microorganisms has preceded and forecast an analogous reaction in animal tissues.

There is evidence to suggest that there exists a biologically active form of acetic acid but its identity has not been established. In the present review the terms acetic acid, acetyl and C_2 unit will be employed, sometimes interchangeably, without any intent of specifying a chemically defined entity.

THE ACETYLATION REACTION. The acetylation of foreign amines and amino acids will be discussed in some detail because this reaction has been of great value for the elucidation of various aspects of acetic acid metabolism in animal tissues.

Acetylation is one of the processes occurring with foreign amines prior to their elimination from the animal body. To what extent acetylation takes place in a given case will depend on the chemical nature and the dosage (6) of the foreign substance and on the species of the test animal. In rats, sulfanilamide is in part excreted unchanged, partly converted to N-acetyl sulfanilamide and oxidized to a hydroxyl compound which can subsequently be conjugated with glucuronic acid. No acetylation of aromatic amines such as sulfanilamide occurs in the dog (7, 13). Also, no unique mechanism exists for the disposal of foreign α-amino acids of the type represented by γ-phenyl α-aminobutyric acid or cyclohexyl alanine. Depending on experimental conditions, the feeding of these amino acids will yield varying proportions of N-acetyl amino acid, hydroxy acid, keto acid and oxidation products derived from the latter (8). Reactions of this kind have been classified as detoxications, a pharmacological term which is generally taken to mean that the body tends to modify foreign substances in such a manner as to render them less injurious to the cell either by increasing their solubility in biological fluids, thus preventing deposition of foreign bodies, or by conversion to derivatives which have less affinity for enzyme systems. The observation that in some instances the toxicity and insolubility of the conversion product exceeds that of the original compound, as in the case of acetylsulfanilamide (9), has made it evident that this type of change undergone by the foreign substance in animal tissues could no longer be adequately described as a detoxication (10, 11). Available evidence indicates that there are but few organic compounds which are entirely refractory to attack by animal cells. The fact that substances are metabolized which are not normally constituents of animal tissues does not necessarily imply the existence of a multitude of enzymes which operate only when the foreign substance is offered to the cell. Thus the coupling of glucuronic acid not only with camphor, borneol or phenol, but also with normally occurring steroid alcohols, or the acetylation of natural as well as of foreign amino acids suggests that these conjugation reactions are not induced by the administration of foreign substances but merely reflect events which are part of the normal intermediary metabolism. The acetylation of a foreign amine may then be attributed to an enzyme which has as its normal function the acetylation of a naturally occurring amine. The enzymes concerned in "detoxication" reactions have not been sufficiently characterized to answer the question whether there exist normal substrates for the enzymes which act upon foreign substances. The enzyme which catalyzes the acetylation of sulfanilamide has been purified (2), but the investigation of substrate specificity has not been extended to normally occurring amino acids. p-amino benzoic acid is acetylated by the same system although at a somewhat slower rate than sulfanilamide.

As an example of a purified enzyme which does not differentiate between a normal and foreign substrate, the l-amino acid oxidase of Green et al. (12) may be mentioned. In the presence of the enzyme, phenylaminobutyric acid, a foreign amino acid, is oxidatively deaminated more rapidly than any of the natural amino acids. Evidently, in this case enzymatic activity is determined

by the α amino acid configuration and not by whether the amino acid is foreign to the cell.

The view that metabolic occurrences with foreign substances in animal tissues mirror normal events formed the basis for the classical investigations of Neubauer, Knoop and Dakin, who employed phenyl-substituted compounds as models for the study of fatty acids and amino acid metabolism. It attests to the validity of the original concepts that the conclusions arrived at by these investigators some forty years ago have not required any fundamental revision.

Acetylation of amines. Following the early finding of Cohn (13) that p-nitrobenzaldehyde was converted by the dog to acetyl p-aminobenzoic acid, acetylation was more closely investigated with the phenyl derivatives of glycine and of α-aminobutyric acid (8, 14). Although these experiments were designed to study synthesis and degradation of amino acids, they yielded equally important evidence with regard to the participation of acetic acid in intermediary metabolism. Knoop (8) showed that the feeding of racemic phenylaminobutyric acid led to the excretion of an optically active acetyl derivative which he found to be identical with the acetyl compound formed in the animal from γ-phenyl α-ketobutyric acid. Knoop erroneously assigned the unnatural configuration to this acetyl amino acid and concluded that the acid of natural configuration had been converted to the acetyl derivative of the unnatural isomer. Similarly the acetyl derivative of what Knoop believed to be l-phenylalanine was excreted unchanged while the acetyl derivative of the amino acid of opposite configuration, judged from the behavior of the racemic acetyl amino aid, appeared to be readily metabolized (15). In view of these findings Knoop was unable to retain his original suggestion that acetylation was a process of normal amino acid metabolism. However, when the configuration of the stereoisomers of phenylalanine was established by du Vigneaud and Meyer (16) it became clear that the configurations assigned by Knoop were incorrect and that, as was subsequently shown by du Vigneaud and Irish (17), the acetylphenylalanine which was metabolically inert belonged to the d-series. The same authors demonstrated that the acetyl compound excreted after feeding of racemic phenylaminobutyric acid was derived from the l-acid and that it was the amino acid of unnatural configuration which underwent optical inversion. These findings re-established the experimental basis for the acetylation hypothesis. Knoop had originally suggested that acetylation was part of a synthetic process leading to acetylamino acids by condensation of the corresponding keto acids with pyruvic acid and ammonia, in analogy to the in vitro reactions described by Erlenmeyer (18) and deJong (19). That acetylation was not necessarily linked to the amination of keto acids was shown by du Vigneaud et al. (20), who found that phenylamino butyric acid labelled with N^{15} retained during its conversion to the N-acetyl derivative nearly all of the N^{15} originally present, indicating that acetylation of the amino or imino acid had occurred. On the other hand, N^{15} was entirely replaced by normal nitrogen when similarly labeled d-phenylaminobutyric was converted to the acetyl l-acid, a result which pointed to the intermediate formation of phenylketobutyric acid. No need exists to postulate two separate

mechanisms for the acetylation of the isomeric amino acids if it is assumed that acetylation is preceded by inversion of the d- to the l-acid. The same authors also studied the acetylation of phenylaminobutyric acid in animals whose body fluids were enriched with D_2O. One atom of deuterium was introduced at the α carbon atom during the acetylation of both isomeric phenylaminobutyric acids. This can be explained readily in the case of the d-acid as resulting from the reductive amination of the intermediate keto acid by hydrogen derived from the labeled body fluids. On the other hand, introduction of deuterium at the α carbon atom of the acetyl l-acid without simultaneous detachment of nitrogen, as shown by the retention of N^{15}, can be visualized only if there exists an equilibrium between amino and imino acid which permits reversible hydrogenation and dehydrogenation. du Vigneaud et al. (20) therefore suggested a mechanism of acetylation in which condensation of the imino acid with pyruvic acid was the initial step, a scheme akin to that of Knoop (8) with the difference that the imino acid takes the place of keto acid plus ammonia. The formulation of the reaction was based on the assumption that pyruvic acid was the most likely source of acetyl. The subsequent demonstration that acetic acid is an acetylating agent for amines (1, 21) eliminated the necessity to associate the reversible hydrogenation-dehydrogenation, as indicated by the uptake of isotopic hydrogen at the α carbon of acetyl amino acids, with the acetylation reaction. It is more likely that amino-imino acid equilibria are established independently of the acetylation reaction as a general property of amino acids. In some instances the acetylation of foreign amino acids of unnatural configuration seems to proceed more rapidly than oxidative deamination and subsequent inversion. S-benzyl d-cysteine (22), p-bromophenyl- and p-bromobenzyl d-cysteine (23) are in part converted to acetyl d-amino acids. These results provide additional evidence that acetylation can be independent of the amination of keto acids.

Precursors of acetyl groups. From the many attempts to identify the nature of the acetyl precursor by balance experiments with intact animals no clear picture has emerged. For instance, there are reports both that the dietary addition of acetate depresses (24, 25) and increases (26) acetylation of sulfanilamide. In balance experiments with intact animals a significant answer is to be expected only if the exogenous supply of precursor is the limiting factor. The probable acetyl precursors such as acetate, pyruvate, or acetoacetate must be available from metabolic reactions in much greater quantities than are required for the acetylation of the relatively small amounts of foreign amine which can be given to animals so that the dietary addition of precursors is unlikely to influence the yield of acetylated product. Direct evidence concerning the sources of acetyl groups for foreign amines has been obtained from feeding experiments with isotopically labeled test substances. Bernhard has administered to humans and rabbits deuterio acetic acid (21) and deuterio ethanol (27), as a source of acetyl for either sulfanilamide, p-amino benzoic acid or cyclohexylalanine (28). In all cases high concentrations of deuterium were found in the excreted acetyl derivatives, demonstrating the conversion of both acetate and ethanol to acetyl groups for aromatic amines and foreign α amino acids. Similar

results were obtained with the stereoisomers of phenylaminobutyric acid by Bloch and Rittenberg (29, 30), who studied the quantitative aspects of the acetylation reaction. When equal amounts of labeled acetate are fed together with either d- or l-phenylaminobutyric acid, or with sulfanilamide or p-aminobenzoic acid, the excreted acetyl groups contain in all cases about the same concentrations of isotope. This indicates that acetic acid is an equally efficient source of acetyl groups for the two types of foreign amines represented by phenylaminobutyric acid (α amino acid) and sulfanilamide (aromatic amine). From the data obtained it was concluded that acetic acid is the main source of acyl groups in the acetylation of foreign amines. The level of isotope in the excreted acetyl groups amounts to a relatively small fraction only of the labeled acetic acid administered. This dilution of isotope is believed to result from the mixing of the exogenous labeled acetate with acetic acid arising in intermediary metabolism (30).

The ability of acetic acid to acetylate amines is also clearly indicated by experiments with surviving tissues and tissue extracts. Acetylation of sulfanilamide has been investigated in slices of guinea-pig liver by Klein and Harris (1) and in homogenate and extracts of pigeon liver by Lipmann (2). In these investigations the aromatic amines, sulfanilamide and p-aminobenzoic acid, were used as acetyl acceptors because conjugation could be followed conveniently by determination of the disappearance of diazotizable amine. That acetyl sulfanilamide was the product formed was established by isolation. In these systems acetate was the most potent acetyl donor, while acetoacetate and pyruvate showed a lesser effect. The results suggest that the immediate reactant is either acetic acid itself or a C_2 compound which under these conditions arises more readily from acetate than from pyruvate or acetoacetate. Five times as much acetylsulfanilamide was formed under aerobic than under anaerobic conditions (2). The linking of acetylation with an oxidative process is further illustrated by the finding that the acetate effect is abolished by oxidation poisons such as cyanide or iodoacetate (1). Although oxygen does not enter into the formal equation of the acetylation by acetate, aerobiosis is required to supply the energy for the condensation. Lipmann has demonstrated that acetylation of sulfanilamide takes place also anaerobically if "energy rich" phosphate in the form of adenosine triphosphate is provided (2). It is not likely that the function of adenosine triphosphate consists in the formation of acetylphosphate since the addition of this compound does not elevate the level of anaerobic sulfanilamide acetylation. The enzyme system concerned in the acetylation of these aromatic amines is most abundant in pigeon liver and has been found to be extractable (2).

A different enzyme seems to be responsible for the acetylation of α amino acids by acetate. Bloch and Borek (31) found, with the aid of deuterio acetate, acetylation not only of l-phenylaminobutyric acid but also of the natural amino acids, leucine and phenylalanine. The yields of acetylamino acids were much higher with slices from rat liver than from those of guinea pig or pigeon. After incubation with labeled acetate, amino acid and non-isotopic acetylamino acid as a carrier, the quantities of acetylamino acid formed could be calculated from

the isotope content of the isolated acetylamino acid. The amounts of acetyl compound obtained with phenylaminobutyric acid were much greater than with the two natural amino acids, presumably because the acetyl derivative of the foreign amino acid is not further metabolized while the acetylated natural amino acids seem to be short-lived intermediates. Thus, under the same experimental conditions the acetylderivative of a natural amino acid, acetylglycine, was found to be deacetylated.

These results indicate that aromatic amines and both natural and foreign α-amino acids effectively utilize acetic acid as a source of acetyl groups. It is evident that whatever the nature of the C_2 compound which is the immediate acetyl precursor, it can readily be formed from acetic acid itself. The effect of adenosine triphosphate on the acetylation reaction observed by Lipmann (2) suggests that the condensation between the carboxyl group of acetic acid and the amine involves the elimination of phosphoric acid rather than of water. Since the effect of adenosine triphosphate cannot be attributed to the intermediary formation of acetylphosphate, the possibility remains to be explored that phosphorylated sulfanilamide or an acetyl derivative of adenosine triphosphate may be formed as intermediates.

The existence of a second mechanism of acetylation and of another source of acetyl groups is indicated by results obtained with labeled alanine (29, 30). Alanine which contains deuterium at the α and β carbon atoms causes the excretion of deuterio acetyl groups when phenylamino butyric acid is simultaneously administered, but only a slight excess of isotope is found in the acetylation product when sulfanilamide or p-aminobenzoic acid are fed as acetyl acceptors. Since alanine, in contrast to acetate, provides acetyl groups for α amino acids ony, its effect cannot be explained on the basis of intermediate formation of acetic acid. The effect of alanine may reasonably be ascribed to the condensation of a C_3 compound, presumably pyruvic acid, with an imino acid and subsequent decarboxylation of the condensation product to an N-acetyl derivative. In this case the failure of alanine to acetylate aromatic amines can be readily understood since imino acids cannot arise from aromatic amines.

The above result which provides experimental evidence for the acetylation schemes suggested by Knoop (8) and du Vigneaud et al. (20) is of interest in several respects. The reaction:

$$CH_3COCOOH + RCH(NH_2)COOH + \tfrac{1}{2}O_2 \rightarrow RCH(NHCOCH_3)COOH + CO_2 + H_2O \quad [1]$$

can be assumed to proceed with an overall decrease of free energy so that it becomes unnecessary to postulate the intermediate formation of an active C_2 compound as the immediate precursor of N-acetyl groups. This is in contrast to the acetylation of amines by acetic acid which is an endergonic process and therefore requires coupling with an energy yielding reaction. The conversion of pyruvic acid to acetyl amino acid, by effecting the degradation of a three carbon compound to a C_2 unit without intermediate formation of free acetic acid, offers a possible mechanism for the conversion of pyruvate to the oxidation level of

acetate, a step in the oxidation of carbohydrate which is little understood at present.

That there may exist a connection between acetylation reaction [1] and the formation of acetoacetate from pyruvate has been pointed out by Krebs (32). The fact that ketone body formation from pyruvate in liver is greatly enhanced in the presence of ammonia (33, 34) can be explained on the basis of acetylamino acid formation from keto acid, ammonia and pyruvate, and subsequent condensation of acetyl groups to acetoacetate Formation of an N-acetyl derivative from pyruvate has so far been demonstrated only with phenylamino-butyric acid but not with natural amino acids. Also, the ability of N-acetyl groups to yield acetoacetate remains to be tested.

The direct participation of acetoacetate as an intermediate in the acetylation of amines by acetic acid has been considered but appears unlikely in view of the following findings (29, 30): deuterioacetyl groups are formed from butyric acid labeled with heavy hydrogen either at the α or the γ carbon atoms, indicating that both two carbon moieties of the intermediate acetoacetic acid are employed for acetylation. Secondly, deuterioacetic acid which also contains C^{13} at the carboxyl position affords N-acetyl groups in which the relative proportion of C^{13} and D remains unchanged. The formation of acetoacetate from such doubly labeled acetate would entail replacement of deuterium by normal hydrogen, but no loss of C^{13}, as the result of the following reactions:

$$2CD_3C^{13}OOH \underset{+H_2O}{\overset{-HDO}{\rightleftarrows}}$$

$$CD_3C^{13}OCD_2C^{13}OOH \underset{+H}{\overset{-D}{\rightleftarrows}} CD_3C^{13}ONa{=}CDC^{13}OOH \qquad [2]$$

$$\downarrow$$

$$(CD_3C^{13}O) + (CH_3C^{13}O)$$

Any acetyl formed by way of acetoacetate should therefore have a diminished $D:C^{13}$ ratio. Since the conversion of acetate to N-acetyl is not accompanied by any loss of carbon-bound deuterium, acetoacetate may be ruled out as an intermediate step. The fact that the $D:C^{13}$ ratio is not affected in a reaction in which acetic acid is utilized also speaks against a rapidly occurring interconversion of acetate and acetoacetate in the intact animal. The significance of this finding with regard to fatty acid catabolism will be discussed later.

Metabolism of acetylamino acids. The occurrence of a process by which acetyl groups can be distributed between amino acids is suggested from data obtained with labeled acetylamino acids (35). These experiments were carried out in order to determine whether the acetyl derivatives of natural amino acids can be biologically deacetylated. Any acetic acid liberated from these compounds should become demonstrable by its capacity to acetylate foreign amines. When the deuterio acetylderivatives of glycine, alanine, leucine or glutamic acid were fed together with phenylaminobutyric acid, deuterium was found in the excreted

acetyl phenylaminobutyric acid, but the level of isotope was very much higher than that obtained after feeding of equivalent quantities of labeled acetic acid. This result was interpreted as showing that the foreign amino acid had received acetyl by direct transfer from the administered acetylamino acid and that in the process no intermediate formation of free acetic acid had occurred. The high isotope concentrations in excreted acetylamine were not obtained when p-amino benzoic acid was the foreign amine. In the transfer of acetyl groups an α amino acid is therefore necessary as an acetyl acceptor. Only the acetyl derivatives of natural amino acids can serve as acetyl donors; the labeled acetyl derivatives of d-alanine, d-leucine and sarcosine failed to yield deuterio acetyl groups. These results paralled those obtained with acetyl d-amino acids in growth experiments and substantiate the conclusion that no biological mechanism is available for the deacylation of acetyl d-amino acids. The fact that α-amino acids are necessary both as acetyl acceptors and as acetyl donors may perhaps indicate that the transfer reaction is coupled with dehydrogenation at the α carbon atom of the amino acids and that in order to allow the reaction to occur the steric configuration of the hydrogen at the α carbon must be that of the natural amino acids. Data on the acetyl transfer reaction have so far been obtained only with phenylaminobutyric acid as acetyl acceptor. *In vivo* the acetyl derivatives of natural amino acids do not accumulate sufficiently to permit their isolation. Transacetylation as a process which operates between two natural amino acids therefore remains to be demonstrated.

Organic chemical reactions representing intermolecular transfer of acetyl groups to amino acids have been known from the investigations of Bergmann et al. Bergmann, du Vigneaud and Zervas (36) found that the acetyl groups of diacetyl diketopiperazine are readily shifted to free amino acids. The reaction proceeds in aqueous solution at room temperature:

Acetyl anhydropeptide + amino acid $\rightarrow$ anhydropeptide + acetylamino acid [3]

Acetyl transfer was also shown to occur with a histidine derivative which contained the acetyl substituent at one of the imidazole nitrogen atoms (37). Evidently an acetyl group is more reactive when linked to a secondary nitrogen and its transfer to a primary amino group *in vitro* will therefore occur spontaneously. This raises the possibility that the acetyl donor concerned in the biological transfer of acetyl groups may be an N-acetylpeptide rather than a free acetylamino acid. On the basis of these model reactions Bergmann and Zervas (37) suggested that acetylamino acids might play a rôle in the biological formation of peptide bonds. This hypothesis assumes added significance in view of the demonstration that the acetyl derivatives of natural amino acids are biologically formed (31) and that a mechanism exists for the shifting of acetyl groups from one amino acid to another (35).

The formation of an acetylamino acid presumably requires a very similar quantity of energy as the formation of a peptide bond, i.e., 2.5 to 3.0 K cal. (38). This energy can readily be supplied by the acetylation mechanism [1] which involves the conversion of pyruvic acid to N-acetyl. On the other hand, it is

not clear which energy yielding step facilitates the acetylation of amino acids by acetate.

Acetyl amino acids may react with free amino acids in two ways: 1, by acetolysis:

$$\begin{array}{l} R_1CHNH\boxed{COCH_3 \quad HO}OC\cdot CHR_2 \\ \;\;| \qquad\qquad\qquad\qquad\quad\; | \\ \;COOH \qquad + \qquad\quad NH_2 \end{array} \rightarrow \begin{array}{l} R_1CH\cdot NHCOCHR_2 \\ \;\;| \qquad\qquad\quad\;\; | \\ \;COOH \qquad\; NH_2 \end{array} \qquad [4]$$

This reaction should involve only a small change of free energy.

Or 2,

$$\begin{array}{l} R_1CHCOOH \\ \;\;| \\ \;NHCOCH_3 \end{array} + \begin{array}{l} H_2NCHR_2 \\ \;\;\;\;| \\ \;\;\;COOH \end{array} \longrightarrow \begin{array}{l} R_1CHCONHCHR_2 \\ \;\;| \qquad\qquad\;\; \backslash \\ \;NHCOCH_3 \;\; COOH \end{array} \qquad [5]$$

The second reaction, which would yield an acetylpeptide, is analogous to the acetylation of amino acids by acetic acid, an acetamino acid being the acetylating agent instead of acetic acid. It may be pointed out that as a result of acetylation, the amphoteric amino acids are converted into strong acids (39).

The ability of the animal organism to utilize acetylamino acids is also illustrated by the finding that, with the exception of α-N-acetyl lysine (40), the acetyl derivatives of essential amino acids are nutritionally equivalent to the free amino acids. This has been tested in growth experiments with the acetyl derivatives of tryptophane (41, 42), phenylalanine (43), and histidine (44). The ability of acetylglycine to raise the level of hippuric acid excretion in rabbits above the normal suggests conversion of the acetylderivative to the free amino acid (45). In those instances in which the acetyl amino acid can take the place of an indispensable amino acid, the corresponding α keto or α hydroxy acids will likewise be capable of supporting growth (46). It is therefore not clear whether regeneration of the free amino acid from the acetyl derivative involves direct deacetylation by hydrolysis or whether it is due to a process which regenerates the amino acid by way of the α-keto or α-hydroxy acid. The observation that α-N-acetyllysine cannot support the growth of rats on a lysine deficient diet has been interpreted by Neuberger and Sanger (40) as suggesting rapid oxidation of the ϵ amino group of lysine prior to deacetylation. However, it is also possible that the α-keto acid is formed directly without passing through free lysine as an intermediate. The α-keto acid of lysine is not known, but it may be presumed to be incapable of substituting for lysine in supporting growth because the α-nitrogen of lysine, in contrast to that of other amino acids, cannot be replaced metabolically by nitrogen from other sources (47) and because d-lysine cannot support growth (48). On the other hand, Neuberger and Sanger (40) found ϵ-N-acetyllysine to be nutritionally equivalent to lysine, suggesting a process of metabolism in which the acetyl group is removed by hydrolysis to regenerate free lysine. In agreement with this observation Bloch and Rittenberg found deuterio ϵ-N-acetyllysine to yield acetyl groups for the acetylation of foreign amines with an isotope content to be expected from hydrolysis of the acetyl amino acid (unpublished experiments).

The enzyme systems catalyzing deacetylations show specificity with respect to the spatial configuration of the acetyl amino acids. Thus the acetyl derivatives of d-tryptophane (42, 49), d-phenyl alanine (17) and d-leucine (35) are metabolically inert; acetyl d-phenylalanine has been recovered unchanged from the urine of the experimental animals (17). The sulfur of acetyl d-cystine is resistant to oxidation while acetyl l-cystine is readily metabolized (50). The unavailability of acetyl d-amino acids must be due to the inability of animal tissues to deacetylate because the corresponding free d-amino acids can be utilized and can replace nutritionally their optical antipodes (49, 51, 52). This is further illustrated by the failure of the deuterio acetyl derivatives of d-alanine and d-leucine to yield labeled acetyl for the acetylation of foreign amines (35). Some significance may be attached to the fact that whereas deamination occurs readily with both stereoisomers of most amino acids, only the acetyl amino acids of natural configuration are metabolized by animal tissues.

Acetylation may be associated with the formation of amino acids from keto acids and ammonia but it can also occur subsequent to amino acid synthesis. It is thus conceivable that acetylation may be concerned in the catabolism of amino acids. The following findings are of interest in this connection. The observation already mentioned of Neuberger and Sanger (40) that α-N-acetyl-lysine cannot replace lysine for support of growth indicates that the compound is not deacetylated but that it may be directly converted to the nutritionally ineffective α keto acid. Secondly, it has been found that on feeding of acetyl-l-phenylalanine to rats, phenylacetic acid can be isolated from the urine in the form of phenaceturic acid. (Unpublished observation of the author.) Since this compound was not obtainable after the feeding of free l-phenylalanine it would appear that the acetyl derivative was deaminated more readily than the free amino acid.

FORMATION OF ACETIC ACID. The first suggestions of a formation of acetic acid in quantity in animal metabolism were contained in experiments of Knoop (53) and Dakin (54) on the biological degradation of phenyl-substituted fatty acids. The excretion of either phenylacetic acid or benzoic acids (or their conjugation products with glycine), depending on the number of carbon atoms in the aliphatic chain of the test substance, was most reasonably explained as resulting from the successive removals of two carbon fragments, by hydrolysis of an intermediate β keto acid. The exact nature of the two carbon fragments could not be ascertained, but chemical considerations pointed to acetic acid or an acetyl compound. Acetic acid has been proposed repeatedly as an intermediate in the main path of carbohydrate oxidation. Thunberg (55) suggested that oxidative decarboxylation of pyruvic acid gave rise to acetic acid, two molecules of which were dehydrogenated to form succinic acid. In the course of their metabolism, several amino acids form intermediates which are identical with products of either carbohydrate or fatty acid breakdown, pointing to protein as a potential source of acetic acid. Thus it was conceivable that the metabolic paths of the three major dietary and tissue components, fat, carbohydrate and protein, would converge at the two carbon stage and that, once this

stage had been reached, only one common mechanism of oxidation would be required. It is interesting to note that the principal aspects of this theory of intermediary metabolism which, at the time, rested on little experimental evidence, have required revision in a few details only. Experimental evidence in its favor continues to accumulate.

The occurrence of acetic acid in animal tissues. As an intermediate in the main chain of metabolic reactions, acetic acid should continually arise in large quantities. Analysis of tissues and body fluids reveals that in the equilibrium state, acetic acid is present in quantities which are barely detectable (56). Acetic acid in the form of its dinitrophenylhydrazide has been isolated from large quantities of beef liver (57), but it is not clear whether the material thus obtained was an artefact formed during the isolation procedure. Specific methods for the determination of small quantities of acetic acid in biological mixtures are not available; it may also escape detection because it seems to be firmly bound to protein in biological media (58, 59). Nevertheless, even if acetic acid were found to be practically absent, the suggested rôle of acetic acid as a major product of catabolism would not be contradicted. The condition which has to be satisfied, and which applies to all metabolites which lie on the main metabolic paths, is a capacity of the tissue to metabolize the intermediate as rapidly as it is formed. The concentration of the metabolite at any one time will then be irrelevant. The insignificant concentration of acetic acid and other fatty acids of intermediate size in animal tissues has been cited as evidence against the theory of β oxidation (60). If this objection were valid, the same reasoning would apply to numerous commonly accepted intermediates of protein and carbohydrate metabolism, because the demonstration of their presence in normal tissues has not been experimentally feasible.

Evidence for the occurrence of a metabolite which does not normally accumulate in sufficient amounts to permit analytical detection can be obtained with the aid of the isotope dilution method. If the isotopic analogue of a suspected intermediate is administered to an animal, or added to an isolated system, and subsequently recovered either as such or in combined form, a diminished isotope content of the isolated material will indicate that the test substance had merged with identical unlabeled molecules formed in the tissues. Bernhard, in an investigation of the acetylation of foreign amines with labeled test substances (21) found that the acetyl groups of excreted acetyl sulfanilamide contained 5 per cent to 20 per cent as much isotope as the deuterio acetic acid added to the diet. He concluded that the fraction of acetyl groups contributed by exogenous acetate was related to the difference of isotope content in the test substance and excreted acetyl. These findings suggested that the acetyl precursor consisted of a mixture of dietary and endogenous acetic acid. The possibility that the acetic acid excreted by the animal in the form of acetylamine was a representative sample of the acetate existing in the metabolic pool offered an experimental approach to a quantitative determination of this metabolite in the intact animal. From such experiments, Bloch and Rittenberg (30) have concluded that the isotope dilution obtaining in the acetylation process can be attributed to the merging of labeled

dietary acetate with acetate arising in metabolism. The magnitude of the dilution factor indicates that the quantities of acetic acid produced are large.

In order to ascertain the correctness of this conclusion it was necessary to consider the following sources of isotope dilution apart from that caused by endogenous acetate: 1, rate of the acetylation reaction; 2, labilization of the C—H bond in the methyl group of deuterio acetic acid; 3, acetylation by precursors other than acetic acid.

1. Since the acetyl derivatives of foreign amines are not stored by the animal but excreted into the urine soon after their formation, the rate of the reaction need not be considered. The isotope concentration of excreted acetyl is the same irrespective of the length of time during which labeled acetate is administered. 2. In one of the acetylation experiments, a preparation of acetic acid which contained deuterum in the methyl group and a carboxyl group labeled by C^{13} was administered simultaneously with phenylaminobutyric acid. The acetyl group of the excreted acylamino acid contained the two isotopic markers in much smaller concentrations, but the ratio of C^{13}:D had remained unchanged. This finding eliminates loss of deuterium as a source of isotope dilution and also attests to the biological stability of the C—H bonds in the methyl group of acetic acid. 3. If the acetyl groups were derived not only from acetic acid but also from other precursors, such acetyl groups would be non-isotopic and hence dilute the deuterio acetyl groups which originate from labeled acetate. Evidence for the contention that acetate may be the only source of acetyl groups for aromatic amines (sulfanilamide, p-aminobenzoic acid) and the major one in the acetylation of α amino acids has been discussed above. It is based on the finding that in identical feeding experiments, acetic acid effects much higher isotope concentrations in excreted acetyl than any other compound tested. According to Bernhard (27), labeled ethanol is a better source of acetyl than acetic acid, but in the experiments of Bloch and Rittenberg (29) ethanol and acetate were equally effective. On the basis of available evidence it seems likely that alcohol is rapidly oxidized to acetic acid and thus becomes a source of acetyl groups. Still, the possibility exists that a C_2 compound which is readily available from both alcohol and acetate is the immediate acetylating agent.

The finding that alanine is a source of acetyl groups for foreign α amino acids but not for foreign aromatic amines has been attributed to a condensation of pyruvic acid with the imino acid derived from phenylaminobutyric acid. While these results establish the existence of a second source of acetyl groups for α amino acids, they do not invalidate the contention that in the acetylation of aromatic amines acetic acid serves as the sole source of acetyl groups. It was therefore reasonably well established that in feeding experiments with labeled acetate the change in isotope concentration occurring in the acetylation reaction provides a direct measure of endogenous acetate production. By applying the equation for isotope dilution (61), the quantity of acetate formed by the rat in a 24-hour period was calculated to be 15 to 20 mM per 100 grams, or approximately one per cent of the body weight. This value could be in error in either direction if the ingested acetic acid failed to mix adequately with the acetate

of the metabolic pool, in which case either the dietary material or the endogenous acetate might be employed preferentially for acetylation. However, the following considerations lend support to the correctness of the above calculation: the isotope concentration of excreted acetyl was found to be directly proportional to the total quantity of labeled acetate administered per unit of body weight, and hence the calculated value for acetate production in vivo is independent of dosage variation over a wide range. This finding points to a rapid merging of dietary acetate and the acetic acid formed in the tissues.

In vitro studies by Klein and Harris (1) indicate that the acetylation reaction is confined to the liver. The sample of acetate which becomes trapped by forming acetylamine is therefore representative only of the hepatic acetate pool. Of the acetic acid of extrahepatic origin, only the fraction which escapes oxidation at the site of formation and is dispatched to the liver will be measured in the acetylation reaction. The resultant error cannot be assessed at the moment, because it is not known to what extent acetyl-yielding reactions take place in extrahepatic tissues.

Lorber et al. (62) perfused the isolated heart with labeled acetic acid and found that the latter was apparently diluted by normal acetate in amounts corresponding to those calculated for the intact rat (30).

R. Q. measurements of animals in the fasting state (63) and after hepatectomy (64) have been interpreted as indicating fatty acid oxidation in organs other than the liver. Stadie (65) has calculated that in the animal depleted of carbohydrate stores, the energy requirements cannot be met by hepatic production of ketone bodies. The occurrence of extrahepatic fat oxidation is illustrated by experiments of Lehninger, who showed that palmitic and octanoic acids are metabolized by a heart muscle preparation to yield succinate and ketoglutarate (66). An independent estimate of the size of the metabolic acetate pool has been made on the assumption that a molecule of fatty acid of average size (C_{16}) yields on degradation eight molecules of acetate. With the aid of available data on rates of fatty acid regeneration (67) it is possible to calculate how much acetate should arise in a given period of time if fatty acids were the only source. The quantity thus calculated agrees closely with that obtained experimentally by isotope dilution measurements.

Acetic acid formation from fatty acids. Fatty acid oxidation[1] has been investigated in a variety of systems, viz., in the fasting or diabetic animal, in perfused livers, in various liver preparations, and in the intact animal with the aid of labeled test substances. Under all conditions except those mentioned last, acetoacetic acid is the principal product of the oxidation process. Leloir and Muñoz (68) detected acetic acid in the course of octanoate oxidation by liver, but the amount was insignificant compared with the quantities of ketone bodies which accumulated. Deuel and co-workers (69, 70) found a uniform increase of ketone body excretion with increasing chain length of the fatty acid, administered as salt or ethyl ester, to starving rats. In isolated liver, the yields of

[1] For a detailed consideration of this subject the reader is referred to the review by Stadie (60).

ketone bodies per molecule of fatty acid obtained by different investigators differ widely (71 to 74), depending apparently on the previous treatment of the liver tissue. *In vitro*, fatty acids appear to become increasingly resistant to oxidation as the length of the chain increases. The depressing effect of higher fatty acids on the respiration of isolated tissue, noted by Quastel and Wheatley (71), has been attributed to their surface activity. Lehninger has been able to show, however, that oxygen uptake can be restored and that the higher fatty acids can be oxidized by addition of adenosine triphosphate to the liver system (74). Under the experimental conditions of this investigator the carbon atoms of some even and odd numbered fatty acids (C_8, C_7) could be accounted for as acetoacetate in nearly quantitative yields.

Since acetoacetic acid accumulates as the principal product of fatty acid oxidation, it is not surprising that acetoacetate rather than acetate has frequently been emphasized as the primary oxidation product. On the other hand, recent experiments have furnished clear evidence for the correctness of the principle of β oxidation as proposed by Knoop (53). Stepwise degradation of a biologically occurring fatty acid by elimination of two carbon atoms was first demonstrated by Schoenheimer and Rittenberg (75). Palmitic acid isolated from the tissue fat of animals which had received deuterio stearic acid had a sufficiently high isotope content to show that it arose directly from the C_{18} acid by shortening of the chain. Their data indicated that myristic acid had been similarly formed by subsequent elimination of two carbon atoms from palmitic acid. The further degradation could not be studied because tissue fat contains only traces of the acids with less than 14 carbon atoms.

The two carbon fragment which is detached from the fatty acid chain has not been identified as such, although its intermediate formation has been clearly established. Weinhouse, Medes and Floyd (76) have investigated the oxidation of octanoic acid which was labeled at the carboxyl group with C^{13}, in liver slices. The acetoacetic acid which accumulated as the main product, contained labeled carbon in nearly equal concentrations at the carbonyl (β carbon) and carboxyl positions. Acetoacetate with such isotope distribution could not have resulted from the splitting of the C_8 acid into two four carbon compounds (multiple alternate oxidation), nor could it have been formed exclusively from carbon atoms 5 to 8 of the octanoic acid (classical β oxidation). Acetoacetate must have been formed, at least in part, by random condensation of C_2 fragments.

The ability of fatty acids to yield acetyl groups in the intact animal can be tested by exploiting the fact that foreign amines are converted in vivo into metabolically inert acetyl derivatives. Since acetic acid is an effective source of such acetyl groups, the assumption can be made that the appearance of labeled acetyl amine after feeding of labeled precursor is a measure of acetyl formation. The following deuterio-fatty acids were found by Bloch and Rittenberg (29) to be sources of deuterioacetyl groups: butyric acid, n-valeric acid, isovaleric acid and myristic acid; propionic and undecylic acids failed to yield acetyl. The results obtained were found to conform with the assumption that β oxidation and removal of acetyl groups from the β keto acids was the principal event in the

metabolism of the fatty acids. It is clear that deuterium is not always an adequate label for the study of carbon chain interconversions, and the results obtained on acetyl formation are in most cases qualitative only. All deuterium attached to the β carbon atoms will be lost by oxidation and further loss of deuterium is to be expected due to keto-enol tautomerism involving the methylene hydrogens adjacent to the keto groups. Thus, α, β dideuteriobutyric acid yields acetyl of much lower isotope concentration than butyrate which contains deuterium at the β and γ positions. If acetoacetate is the intermediate in the butyrate-acetyl conversion, loss of deuterium from the labile α position and therefore a lowered isotope content of the acetyl fragment derived from the carboxyl and α carbon atom of butyric acid would result. The behavior of the two preparations of deuterio butyric acid suggests breakdown of the intermediary β keto acid into two identical acetyl fragments.

The various β keto acids which are formed successively from the higher fatty acids will suffer the same loss of deuterium at the α carbon as the α fragment of butyric acid, so that only the deuterium bound to the terminal carbon remains in stable linkage. In the intact animal, the long chain fatty acids, in contrast to those having a small number of carbon atoms, are only partly catabolized immediately to yield acetyl groups. For example, a large portion of administered myristic acid will be deposited in the tissues as such and will also be converted to palmitic and stearic acids by chain elongation. In this case the deuterium content of excreted acetyl will not be a quantitative measure of the amounts of acetyl produced.

Metabolism of odd numbered fatty acids. The odd numbered straight chain fatty acids are not encountered in animal metabolism but they are oxidized in vivo and in vitro as readily as the corresponding even numbered fatty acids. In isolated liver (72) and in the intact animal (77) the fatty acids possessing an odd number of carbon atoms (except propionic acid) are precursors of ketone bodies although the quantities produced are smaller than those afforded by the neighboring fatty acids with an even number of carbon atoms. Propionic acid is presumably oxidized to a three carbon compound which can be converted quantitatively to carbohydrate. Bloch and Rittenberg have concluded that pyruvic acid cannot be an intermediate in this conversion since deuteriopropionic acid, in contrast to labeled alanine, failed to provide acetyl groups in the acetylation of phenylaminobutyric acid (29). Deuterio n-valeric acid formed labeled acetyl to an extent which indicated that one two-carbon fragment arose per mole of valerate. Cohen found that isovaleric acid is a much more potent source of acetoacetate in liver slices than n-valerate (73). In the intact animal labeled isovalerate yielded more deuterioacetyl than equivalent amounts of the n-valerate (78). From the data obtained, it could not be decided whether the breakdown of this branched chain acid proceeds by way of demethylation to butyric acid or by initial oxidation at the tertiary carbon atom to acetone and a C_2 fragment. Undecylic acid labeled with deuterium at the 10 and 11 positions did not yield deuterio acetyl. β oxidation of this eleven carbon acid may be visualized to lead eventually to a three carbon fragment which contains the

isotopic hydrogen but which, like propionic acid, would not be degraded directly to a C_2 unit. The behavior of the odd numbered fatty acids can be readily understood if it is assumed that their metabolism proceeds in a fashion analogous to that of the even-numbered fatty acids with the difference that their terminal three carbon atoms are, like propionic acid, converted to carbohydrate. From an odd numbered acid having n carbon atoms, only n−3 carbon atoms will then be available for ketone body formation so that less acetoacetate will be formed per molecule than from the neighboring even numbered fatty acid.

The demonstration of the ketogenic effect of n-valeric acid (72, 77) has had an important bearing on the development of more recent theories of fat oxidation. This finding could not be adequately explained by either classical β oxidation or by the hypothesis of multiple alternate oxidation. MacKay et al. (77) suggested that the five carbon chain of valerate was degraded to a three carbon and a two carbon unit and that two of the latter combined to yield acetoacetate. It became evident therefore that ketone bodies could be formed by recondensation of two carbon fragments split off from the fatty acid chain as well as by direct oxidation.

Formation of acetyl from amino acids. It is generally believed that α amino acids are catabolized by oxidative deamination to α keto acids and subsequent oxidative decarboxylation to acids containing one carbon atom less. In this case amino acids which yield saturated fatty acids as intermediates should be sources of acetyl groups. This possibility has been studied with the amino acids leucine and valine which contained carbon chains labeled by deuterium (78). Deuterioleucine yielded labeled acetyl groups to the same extent as isovaleric acid indicating that this acid was formed as an intermediate. The ketogenic effect of leucine (73, 79) may be explained on the same basis. Valine and isobutyric acid both are nonketogenic and are also ineffective as sources of acetyl groups. No information is available concerning the steps by which isobutyric acid is further metabolized except that three of its four carbon atoms are convertible to glycogen (80). It should be of interest to ascertain whether the ketogenic action of phenylalanine and tyrosine is associated with acetyl formation, particularly whether the ring structure of these amino acids can be degraded to two carbon units.

Mechanisms of fatty acid oxidation. A consideration of all pertinent data obtained in experiments with intact animals and with *in vitro* systems leads to the conclusion that successive β oxidation, followed by removal of C_2 fragments, is the principal event in fatty acid oxidation. For a number of years undue emphasis was placed on experimental observations which appeared to contradict β oxidation as originally proposed by Knoop. It should be pointed out that the data which led to a criticism of the theory of β oxidation were obtained with either isolated tissues or starving or diabetic animals. Since under these conditions acetoacetic acid and not acetic acid accumulates, the ketone bodies came to be regarded as primary oxidation products not only of the terminal four carbon atoms but of the entire fatty acid chain. Moreover, the theory of multiple alternate oxidation was advanced as a mechanism for fat oxidation

because, *in vivo* and *in vitro*, the fatty acids containing six and more carbon atoms gave a higher yield of acetoacetic acid than butyric acid, while Knoop's theory provided for the formation of only one mole of keto acid regardless of the length of the chain. However, the formation of ketone bodies in excess of one mole per mole of fatty acid is not at variance with β oxidation if it is ascribed to the recondensation of acetic acid to acetoacetic acid. The occurrence of this reaction was early demonstrated by Loeb (3) and Friedemann (81), who observed acetoacetate formation on perfusing liver with acetate. This finding has been confirmed repeatedly and under a variety of conditions, viz., by Jowett and Quastel and by Leloir and Muñoz in liver slices (72, 68), by Monguio and by MacKay et al. in intact animals (4, 5), and with C^{13} labeled acetate in starving animals by Swendseid et al. (82) and by Medes et al. in liver slices (83). It was recognized by MacKay et al. (5) that the synthetic formation of acetoacetate from acetate provided a basis for a theory of fat oxidation which retained Knoop's basic concept of β oxidation and ascribed the formation of "excess" acetoacetate to a subsequent event, namely, a recombination of two molecules of acetate. The data of Weinhouse et al. with carboxyl labeled octanoic acid (76) demonstrate that acetoacetate can be formed by random coupling of C_2 fragments which are split off from the fatty acid chain. There are, however, two possible pathways which conform with the observed isotope distribution:

$$\text{(a)}\quad CH_3(CH_2)_6\,C^{13}OOH \rightarrow 4C_2 \rightarrow 2CH_3C^{13}OCH_2C^{13}OOH \qquad [6]$$

$$\text{(b)}\quad CH_3(CH_2)_6\,C^{13}OOH \rightarrow C_2^{13} + C_2 + CH_3CH_2CH_2COOH$$

$$\begin{aligned} CH_3CH_2CH_2COOH &\rightarrow CH_3COCH_2COOH \qquad [7] \\ C_2^{13} + C_2^{13} &\rightarrow CH_3C^{13}OCH_2C^{13}COOH \end{aligned}$$

Reaction (a) does not indicate the mechanism which yields the four C_2 fragments from the fatty acid chain; it implies that all of the acetoacetate is of synthetic origin. Reaction (b) is a modified version of classical β oxidation and retains the view that the four terminal carbon atoms are a direct source of acetoacetic acid. In this case the accumulated acetoacetic acid would be a mixture of "primary" and synthetic molecules. In similar experiments with carboxyl labeled butyric acid the resulting acetoacetate contained an unequal distribution of C^{13} at the carboxyl and carbonyl carbon atoms (85). This acetoacetate could therefore not have been formed by random coupling of two equivalent C_2 fragments. Medes et al. (85) have pointed out that their data can be represented by the following equation:

$$CH_3CH_2CH_2COOH \rightarrow CH_3COCH_2COOH \rightleftarrows 2CH_3COOH \qquad [8]$$

In this case rapid interconversion of acetate and acetoacetate has to be postulated. This is true for the synthesis of acetoacetate in liver (84) and kidney (86). The reverse reaction was found by Lehninger (59) to proceed slowly in muscle mince, but it is not clear whether in liver and kidney acetoacetate is split at a sufficient rate to account for the redistribution of isotope in accordance with the above equation.

A similar explanation may apply to the results obtained by Morehouse (87) and by Morehouse and Deuel (88), who fed two preparations of deuterio butyric acid and deuterio caproic acid to starving rats and isolated β-hydroxybutyric acid from the urine. The excreted hydroxybutyrate contained nearly twice as much deuterium when β-γ-dideuteriocaproate was fed as the acid excreted after administration of caproate labeled at the α and β positions. It is evident that a splitting of the hexanoic acid into three two carbon units followed by recondensation to acetoacetic acid should yield hydroxybutyric acid containing the same isotope concentration in both cases. On the other hand, if the excreted β-hydroxybutyric acid was derived mainly from "primary" acetoacetate representing the intact chain of carbon atoms 3 to 6 of caproate, and was formed only in part by recondensation of C_2 units, then β-γ-dideuteriocaproate should yield a hydroxybutyric acid of higher deuterium concentration than α-β-dideuteriocaproate.

Stadie (60) has pointed out that the data obtained by Weinhouse et al. with labeled octanoate are not inconsistent with the theory of multiple alternate oxidation if the random distribution of labeled carbon is the result of a splitting and reformation of acetoacetate. Since it is agreed that the formation of a C_2 fragment is an obligatory step in the breakdown of fatty acids, it remains to be decided whether acetoacetate is the primary oxidation product or whether its formation represents an alternative path for the disposal of C_2 fragments which are initially formed.

Some indications as to the rôle of acetoacetate in fatty acid oxidation are contained in experiments on the acetylation of foreign amines by acetic acid (30). It has been pointed out that the results from these experiments eliminate acetoacetic acid as an intermediate in the conversion of acetate to N-acetyl groups. From the same analytical data it can be concluded[2] that in the intact animal a rapid interconversion of acetate and acetoacetate is unlikely to occur. Doubly labeled acetic acid ($CD_3C^{13}OOH$), when fed and recovered as acetyl phenylaminobutyric acid, retains deuterium and C^{13} in identical relative concentrations. Had acetate been converted to acetoacetate and regenerated, replacement of deuterium by normal hydrogen and loss of deuterium by enolization should have altered the $D:C^{13}$ ratio. The accumulation of acetoacetic acid in fatty acid oxidation in vitro, or in the urine of animals in ketosis may therefore be ascribed to circumstances which interfere with the normal disposal of acetyl or acetate. The isolated liver or liver preparations as ordinarily employed contain little available carbohydrate, and therefore lack the requisite dicarboxylic acids for the oxidation of the intermediates of fatty acid breakdown. Liver tissue appears to have only a limited capacity to oxidize fat beyond the C_2 stage but dispatches the products to extrahepatic tissues for complete combustion. In the isolated liver acetyl will therefore accumulate and become stabilized in the form of acetoacetate. This point is clearly illustrated by the experiments of Lehninger (89) who found that in preparations of washed liver cells, acetoacetate is the

[2] The author is indebted to Prof. D. L. Thomson for having called attention to these implications contained in the experimental data.

sole product of octanoate oxidation, but that on addition of fumarate the quantities of acetoacetate are diminished in favor of α-ketoglutarate, succinate, and citrate. It becomes unnecessary then to assume that acetoacetate lies on the main path of fat oxidation. Acetoacetic acid is utilized under a variety of conditions, but this is probably preceded by its splitting into two carbon units.

Formation of acetyl from pyruvate. There exists at present no conclusive evidence to indicate whether pyruvic acid enters the citric acid cycle by combining directly with oxaloacetic acid or whether this condensation is preceded by a conversion of pyruvate to an intermediate which has the oxidation level of acetic acid. Formation of acetate from pyruvate can be demonstrated to occur in animal tissues, either aerobically (90, 91, 92) or anaerobically by dismutation (90, 91, 93) but this reaction is not believed to represent an important pathway, because under the same conditions acetate itself is oxidized relatively slowly. It is considered more likely that the oxidation product of pyruvate is represented by a two carbon compound which is metabolically more reactive than acetic acid itself. A phosphoroclastic splitting of pyruvate into acetyl-phosphate and formate (or $CO_2 + H_2$) as it occurs in bacteria (94) has not been observed in animal tissues.

According to Barron et al. (95), acetate can be shown to accumulate when liver or kidney slices oxidize pyruvate in the presence of fluoroacetate. According to these investigators, this compound seems to inhibit the oxidative disappearance of acetic acid by competing with acetate for the enzyme. The accumulation of acetate indicates that fluoroacetate may exert its effect by blocking condensations involving the reactive acetyl compound, which may then become stabilized in the form of acetate.

In the presence of malonate, pyruvate oxidation by liver slices or mince affords acetoacetate in quantities up to 25 per cent of the pyruvate metabolized (33, 34, 96). Pyruvate is quantitatively converted to acetoacetate by the washed liver cell preparation of Lehninger (66). When fumarate was added to the same system, acetoacetate formation was suppressed in favor of an increased yield of di- and tri-carboxylic acids. The in vitro conditions which affected the relative yields of ketone bodies and polycarboxylic acids respectively were the same whether octanoic acid or pyruvate served as substrates. On the basis of these findings it would appear that a common oxidation product is formed from both compounds. A view which has been widely accepted is that an acetyl-like two carbon unit forms the intermediate step at which the pathways of carbohydrate and fat metabolism converge.

However, certain difficulties are encountered when an attempt is made to correlate the ketone body formation from pyruvate *in vitro* with the metabolic fate of fat and carbohydrate in the intact organism. It is well known that in the starving or diabetic animal fatty acids enhance ketone body formation, whereas pyruvate or lactate reduce ketosis. Evidently, under these conditions, the two substances do not enter a common path. It is possible that acetoacetate formation from pyruvate is the favored reaction under special conditions in vitro, but it may still be quantitatively insignificant for the pyruvate metabolism of the

intact animal. The hypothesis that a C_2 fragment arises in the oxidation of both carbohydrate and fat would seem to be untenable unless it is assumed that the C_2 fragments are not identical.

The contrasting behavior of labeled fatty acids and pyruvate in the acetylation of foreign amines in intact animals leads to similar conclusions (30). While acetate is an effective source of acetyl for both aromatic amines (sulfanilamide) and α-amino acids (phenylaminobutyric acid), alanine (pyruvic acid) can yield acetyl to foreign α-amino acids only. Moreover, labeled alanine, unlike acetate or butyrate, does not serve as a precursor for cholesterol, either *in vivo* (30) or *in vitro* (97). If pyruvate were converted to the same C_2 fragment which is active in the acetylation of foreign amines and in cholesterol formation, labeled alanine or pyruvate should show the same effect as acetate. Recent experiments by Sonne et al. (98) on uric acid formation contain similar evidence. The feeding of $CH_3C^{13}OOH$ resulted in the incorporation of heavy carbon at carbon atoms 2 and 8 of the purine nucleus. No significant concentrations of C^{13} appeared at these positions when $C^{13}H_3C^{13}HOHCOOH$ was fed, indicating that lactate is not broken down to the two carbon compound employed in uric acid synthesis.

It is conceivable that the differences observed are quantitative rather than qualitative, but it is not likely that the bulk of pyruvate is oxidized by way of the same C_2 fragment which arises in the catabolism of fatty acids. Bloch and Rittenberg have concluded that the amounts of acetate produced in the intact rat can be accounted for fully by the oxidative metabolism of fatty acids and ketogenic amino acids (30). These data also would seem to exclude any considerable breakdown of pyruvic acid by way of acetyl as an intermediate.

Acetylmethylcarbinol (acetoin) has been obtained as a product of the anaerobic metabolism of pyruvate in micro-organisms (99) and in animal tissues (heart muscle) (100). The process is believed to involve oxidative decarboxylation to acetaldehyde, condensation of the latter with a second molecule of pyruvic acid, and subsequent decarboxylation of the condensation product. A scheme involving the following steps: pyruvate $\rightarrow$ acetoin $\rightarrow$ diacetyl $\rightarrow$ acetyl phosphate $\rightarrow$ acetate—has been proposed by Doisy and Westerfeld (24) on the basis of the effect shown by acetoin in the acetylation of p-aminobenzoic acid. The relatively high toxicity of acetoin for animals (101) and the fact that it is in part excreted after reduction to butyleneglycol (102) seems to preclude the possibility that the above scheme represents a major pathway of pyruvate metabolism.

The formation of acetate from pyruvate by an indirect route involving the splitting of intermediates of the tricarboxylic acid cycle has become a possibility in view of the following observations. The fermentation of citrate by yeast (103) and by various bacteria (104, 105, 106) has been found to yield 1 to 2 moles of acetate and smaller amounts of formate and succinate. These results have been interpreted to suggest the following series of reactions: citrate $\rightarrow$ acetate + oxaloacetate $\rightarrow$ pyruvate + CO_2 $\rightarrow$ acetate + formate. The first of these steps, i.e., the reversal of the condensation of acetate and oxaloacetate is believed by Lipton and Barron (107) to be responsible for the positive effect shown by citrate in the acetylation of choline in brain extracts. The reversibility of the initial

condensation reaction in the tricarboxylic acid cycle has not been tested in other systems of animal origin.

Slade and Werkmann (108) have studied the conversion of succinate to acetate by *Aerobacter indologenes* with the aid of C^{13}. Although these authors conclude that their data indicate a splitting of succinate into two molecules of acetic acid, it is equally possible that the acetate arose not from succinate directly but by degradation of either α-ketoglutarate or of a tricarboxylic acid formed from succinate. Irrespective of the pathway involved, these results suggest a possible route for the conversion of carbohydrate intermediates to acetate and fatty acids which circumvents the decarboxylation of pyruvate itself.

Utilization of acetic acid for synthetic reactions. The administration of labeled acetic acid to animals has been found to result in the incorporation of isotope into a variety of tissue constituents: glycogen, cholesterol, fatty acids, the dicarboxylic amino acids, protoporphyrin and uric acid. It thus appears that in the biological formation of the cell constituents acetic acid is of general importance as a source of carbon atoms. It has become increasingly evident in recent years that body constituents of high molecular weight are synthesized by condensation of numerous small sized units rather than by the utilization and rearrangement of preformed large molecules.

In a few cases only the animal organism depends exclusively on an exogenous supply for the maintenance of cellular composition (vitamins, essential amino acids and essential fatty acids). The majority of the tissue constituents can be derived not only from unchanged dietary material but can also arise by synthesis from small molecular breakdown products which are formed in intermediary metabolism. The ability of the animal organism to keep its tissue composition constant under a variety of dietary conditions is an expression of the fact that the required elementary building stones are available from a number of sources.

Older suggestions that cholesterol might be formed either by modification of the preformed steroid structure which is offered to the animal in the form of squalene (109) or plant steroids (110), or by folding of long chain fatty acids (111) have given way to the view that a total synthesis from numerous small units is involved (112). In the synthesis of long chain fatty acids from carbohydrate the sugar molecule is presumably not utilized as such but only after its degradation to C_2 or C_3 fragments. In the formation of the tetra pyrrole structure of porphyrin, two C_2 compounds, acetic acid (30) and glycine (113, 114), serve as a source of carbon and nitrogen and not the preformed five-membered ring structures of proline or pyrrolidonecarboxylic acid (from glutamic acid). The investigations on the biological synthesis of these high-molecular tissue constituents have so far been carried out mainly with intact animals. It is evident that for the synthesis of the more complex molecules a high state of cellular organization will be required. For this reason virtually no attempts have been made to investigate the enzymatic processes concerned in the synthesis of the substances mentioned. It has not been possible to identify any of the intermediate steps in the biosynthesis of fatty acids, cholesterol or porphyrin, nor do indications exist that such intermediates occur normally in measurable quantities.

The utilization of acetic acid for biological syntheses has been demonstrated by the feeding of labeled acetic acid to animals and isotope analysis of the body constituents isolated from the tissues. Insight into the reaction mechanisms has in some cases been gained by determining the location and distribution of isotope at various positions of the synthetic product, but the accessible information is often limited by the lack of suitable methods of degradation. Wood et al. (115) have employed chemical and microbiological degradation procedures for glucose which permit separate isotope analysis of carbon atoms 3 and 6 and of fractions which represent carbon atoms 1 and 6, 2 and 5, and 3 and 4 respectively. Methods suitable for determining the deuterium concentrations at separate positions of deuterioglucose have been described by Stetten and Stetten (116). Although procedures are available for successive degradation of the higher fatty acids they are not applicable to small quantities of biological material. However, analysis of the carboxyl group and of the two moieties obtainable by degradation of oleic acid have yielded valuable information on the mechanism of fatty acid synthesis (117). The distribution of isotope in deuteriocholesterol has been determined by degradation of the steroid molecule into nucleus and side chain (118). No experiments of this type have so far been carried out with protoporphyrin formed biologically in the presence of labeled acetate.

Although it may be feasible to study details of reaction mechanisms by employing test substances which contain more than one isotopic marker, there are certain limitations to the experimentation with intact animals. If, in a synthetic process, acetic acid did not take part as such but in the form of a functional derivative (acetylphosphate, acetylamine) and if the substituent were eliminated in the course of the condensation, isotope analysis of the reaction product may not reveal whether acetic acid reacted as such or in a combined form. On the other hand, experiments with intact animals are indispensable in order to establish the occurrence and the quantitative significance of a process which has been observed in an isolated system.

Glycogen. According to classical concepts, a glycogenic compound is a substance which either increases glycogen deposition in the liver of a fasted animal or enhances glycosuria in the diabetic animal. Acetic acid does not satisfy these criteria, since its feeding neither increases glycogen deposition nor glucose excretion (119). On the other hand, isotope experiments have shown that acetic acid can provide carbon atoms for all positions of the glucose molecule (120, 121). No inconsistency need be seen in these findings, if (a) the meaning of the term precursor is more clearly defined and if (b) it is recognized that the terms glycogenic and ketogenic apply only to special experimental conditions. The revision of classical concepts necessitated by the results obtained with tracer substances has been discussed by Buchanan and Hastings (122) and by Wood (123). If one disregards reactions in which a substance effects an increased formation of a product by stimulation rather than by direct conversion, then two types of precursors have to be considered: substances which afford a net increase of the quantity of product, i.e., reactions in which the quantity of precursor is the limiting factor, and secondly substances which are incorporated either totally or

in part into reaction products but which are without effect on the rate of its formation. Compounds which in classical terminology are classified as either glycogenic or ketogenic belong to the first group while acetic acid with respect to glycogen formation belongs to the second category. Although, as Lorber, Lifson and Wood (121) have shown, acetic acid provides carbon atoms for all positions of the glucose molecule, it probably cannot by itself form hexose (or pyruvate) but only by entering the tricarboxylic acid cycle. Through a series of transformations the tricarboxylic acid yields pyruvate, which is the obligatory intermediate in glucose formation. Acetate utilization for glycogen synthesis is therefore not regulated by the availability of acetic acid but depends on tricarboxylic acid formation and regeneration of pyruvate from the latter. On the other hand the conversion of pyruvate to glycogen does not require the intervention of the citric acid cycle and as a result pyruvate and compounds which are directly convertible to pyruvate will be precursors of glycogen in the sense that they can increase the net amount of glycogen formed. It is evident that the terms glycogenic and ketogenic apply only to conditions in which a change of carbohydrate concentration is observable, i.e., in the starving or diabetic animal. From its behavior in isolated liver, pyruvic acid could be classified as a ketogenic substance, since under such conditions it can be converted in part (124, 33, 90, 96) or quantitatively (125) to acetoacetatic acid. Acetic acid is ketogenic since it increases ketone body excretion in the starving or diabetic animal (5), but as isotope studies have shown, it is also a source of carbon atoms for glucose. The appearance of acetate carbon in all positions of the glucose molecule has been ascribed by Wood (115) to a series of steps by which a tricarboxylic acid formed from acetate and oxalacetate is degraded to pyruvate by way of α-ketoglutarate and the C_4 dicarboxylic acids. If $C^{13}H_3C^{13}OOH$ is employed the resultant pyruvate will be labeled at all three positions but it cannot contain more than half of the isotope concentration present in the acetate which was utilized. On tracing the labeled carbon through tricarboxylic acid, α-ketoglutarate and the dicarboxylic acids to succinate, it becomes clear that only two of the four carbons in the dicarboxylic acids originate from acetate but the isotopic carbon becomes evenly distributed over the entire carbon chain when the asymmetry of the molecule is lost at the succinate stage. Pyruvate will therefore contain C^{13} at all positions. It can be concluded that although acetate carbon appears in all positions of glucose not more than half of the carbon atoms of a glucose molecule can have originated from acetate, the remainder being supplied by oxaloacetate. Acetate per se is therefore incapable of forming glucose.

Carboxylation of acetic acid to pyruvic acid cannot be an important pathway for the utilization of acetate in glycogen formation because in this case the isotope distribution in glycogen would be expected to differ from that actually observed (121).

Fatty acids. McLean and Hoffert (126) have studied lipid formation in yeast from acetate and have concluded from balance data that both steroids and neutral fat could be synthesized directly from acetate without intermediate formation of carbohydrate. In the presence of sulfite steroid synthesis was

inhibited, but fat continued to be synthesized. This suggested that acetaldehyde might be an intermediate in the synthesis of yeast steroids while acetate could condense directly to fatty acids. In their experiments on the metabolism of deuterioacetate in yeast, Sonderhoff and Thomas (127) found the following concentrations of deuterium in the various fractions: carbohydrate 1.6 per cent; fatty acids 14.7 per cent; unsaponifiable 30 per cent, indicating that carbohydrate was not an intermediate in the conversion of acetate to lipids.

Wood et al. (128) have demonstrated the bacterial utilization of acetic acid for the synthesis of lower fatty acids. In the presence of $CH_3C^{13}OOH$, the butyl alcohol formed in the fermentation of *Cl. butylicum* and of *Cl. acetobutylicum* was found to contain C^{13} at carbon atoms 1 and 3. Since these organisms can reduce butyric acid to butanol, the following series of reactions is suggested: 2 acetate $\rightarrow$ butyrate $\rightarrow$ butanol.

The fermentation of alcohol and acetate in *Cl. kluyveri* has been studied by Barker, Kamen and Bornstein (129) The butyric acid formed in the presence of $CH_3C^{14}OOH$ contained approximately equal quantities of heavy carbon in the carboxyl and β positions; in caproic acid equal amounts of C^{14} were presumably present at alternate carbon atoms since the carboxyl group accounted for about one-third of the total isotope content of the hexanoic acid. Important information with respect to the mechanism of fatty acid synthesis in these organisms was obtained by the same authors. Isotopic caproate formed in the presence of $CH_3CH_2CH_2C^{14}OOH$ and normal ethanol contained isotope but this was not present in the carboxyl group. The active C_2 unit derived from ethanol must therefore have condensed with the carboxyl group of butyric acid. Moreover, a breakdown of butyrate to C_2 compounds prior to formation of the C_6 acid cannot have occurred since in this case the carboxyl carbon of caproic acid should have contained the carbon isotope.

Rittenberg and Bloch (117) have fed mice and rats acetic acid which contained C^{13} at the carboxyl position and deuterium in the methyl group. The fatty acids isolated from the tissue lipids contained both isotopes in excess of the concentrations which could have resulted from the oxidation of acetate to $C^{13}O_2$ and D_2O and incorporation of deuterium from the body fluids or C^{13} by a reaction involving CO_2 assimilation. The body fluids of the mice contained 0.09 per cent D and the respiratory CO_2 contained 0.066 per cent C^{13} while the saturated fatty acids isolated from the liver lipids contained 0.42 per cent D and 0.16 per cent C^{13}. As the fatty acids contained both C^{13} and deuterium it is clear that the acetic acid molecule as such had been utilized in the synthetic process. From acetic acid containing isotopic carbon in the carboxyl group, a fatty acid containing C^{13} only at one out of two carbon atoms is to be expected. The incorporation of isotope into the fatty acids could result from the following processes: 1, acetic acid elongates the fatty acid chain by adding at the carboxyl end. In this case the presence of isotopic carbon should be confined to the carboxyl carbon of the fatty acid and decarboxylation should leave a non-isotopic residue; 2, If acetic acid condensed with the ω carbon of the fatty acid decarboxylation should yield CO_2 which contains no excess of isotope; 3, acetate carbon is uniformly

distributed over the fatty acid chain. C^{13} should then be present at the odd-numbered carbon atoms. The CO_2 obtained on decarboxylation of the fatty acid actually contained about twice as high an isotope concentration as the entire molecule, indicating that C^{13} was present at alternate positions, i.e., at the odd-numbered carbon atoms, one of which is the carbon of the carboxyl group. Additional support for the view that the acetic acid carbon was distributed in random fashion over the entire fatty acid chain is afforded by analysis of the degradation products of the "oleic acid" fraction of tissue fat. This consists mainly of oleic acid and some palmitoleic acid and yields on oxidation a monocarboxylic acid, pelargonic acid, derived from carbon atoms 10 to 18 and a dicarboxylic acid (azelaic acid) representing carbon atoms 1 to 9. If a C_{18} fatty acid were formed from nine acetate units then of the nine isotopic carbons, five should be present in the azelaic acid and four in the pelargonic acid fraction, while the deuterium content of the two moieties should be nearly equal. The following isotope concentrations were found: azelaic acid, 0.090 per cent C^{13}, 0.14 per cent D; pelargonic acid, 0.071 per cent C^{13}, 0.15 per cent D. These data suggest a process of fatty acid synthesis involving the multiple condensation of C_2 units. Any scheme proposed as a mechanism of fatty acid synthesis must account for the fact that the fatty acids of animal tissues have an even number of carbon atoms and comprise, if milk fat is included, all members of the series from C_4 to C_{24}. The occurrence of C_2 condensation in the synthesis of fatty acids by animal tissues was first demonstrated by the finding of Stetten and Schoenheimer (130) that deuteriostearic acid is formed from deuteriopalmitic acid by chain elongation.

The reactive two carbon compound undergoing the condensation can be readily formed from acetic acid in the intact animal and in bacterial systems, but its chemical identity has not been determined. Fatty acid synthesis in animal cells has so far not been demonstrated in isolated systems. When rat liver slices were incubated with labeled acetic acid only very small amounts of isotope were incorporated into the fatty acids; significant concentrations of deuterium were found in the liver fatty acids when the buffer medium contained D_2O (131). From this result it cannot be decided whether the uptake of isotope resulted from synthesis of the carbon chain or whether it was due to hydrogenation of unsaturated fatty acids.

The production of C_2 units in the course of fatty acid oxidation and their utilization in fatty acid formation points to a mechanism of synthesis which consists of the reversal of the catabolic steps. This would involve the condensation of two carbon units with acetate and higher fatty acids and the subsequent reduction of β keto compounds to saturated acids. The well established conversion of acetate to acetoacetate may represent the initial step in this process. However, the biological reduction of β-keto acids to the corresponding saturated fatty acids has never been shown to occur.

Experiments with acetic acid containing deuterium as well as C^{13} have thrown some light on the nature of the intermediates involved in fatty acid synthesis (117). When a preparation of acetate was fed in which deuterium and carbon

were present in a ratio of 8:1 the fatty acids isolated contained the two isotopes in a ratio of 2.5:1. Irrespective of the nature of the condensation, there should be introduced for every two deuterium atoms from the methyl group of acetic a minimum of two atoms of ordinary hydrogen from the body fluids. The ratio of $D:C^{13}$ should then become 4:1. The greater loss of D than corresponds to this ratio may have resulted from keto-enol tautomerism of the intermediate β-keto acid. An additional atom of normal hydrogen from the body fluids would also be introduced if the keto acid were converted to the saturated acid by way of the hydroxy- and unsaturated acids.

Although the isotope data suggest that acetic acid can supply the carbon atoms for all positions of the fatty acid chain, it is unlikely that acetate is a precursor in the sense that its administration produces a net increase of synthesized fat. Experimental conditions for studying fat synthesis from specific precursors, similar to those by which the glycogenic or ketogenic properties of a compound can be tested, are not available. Conditions which result in the deposition of excessive quantities of fat in the liver such as choline deficiency do not measure changes in the amount of fat synthesized but reflect deficiencies in fat transport. Stetten and Boxer found that in rats treated with alloxan, fat deposition and incorporation of deuterium from heavy body water are greatly reduced (132). This may be the result of diminished fat synthesis or of a failure to deposit the newly formed fat. The impairment of fatty acid synthesis in this condition is probably not due to a lack of precursor but may reflect, as Stetten and Boxer suggest, a general inability to utilize carbohydrate. Since in the experiment of Rittenberg and Bloch (117) the animals received a fat free diet, carbohydrate and protein must have provided the carbon for fatty acid synthesis. It was estimated that under these conditions about 25 per cent of the carbon atoms of the fatty acids were derived from acetate. While the conversion of carbohydrate to fat is well established from balance experiments, no information is available concerning the intermediate steps. It is generally assumed that synthesis of the fatty acid chains occurs with pyruvate or a two carbon compound derived from pyruvate. It is questionable whether this intermediate is identical with acetic acid because in this case not only one-fourth but all carbon atoms of the fatty acids should be available from the "acetyl" pool. It has been pointed out above that pyruvic acid or alanine are not convertible to the acetyl compound involved in the acetylation of sulfanilamide or in the formation of cholesterol. This has been taken to mean that a direct degradation of pyruvic acid to a C_2 compound does not occur or that the product is not identical with the acetyl formed by fatty acid oxidaton or from acetate itself.[3] In order to allow for the participation of both acetic acid and carbohydrate in fatty acid synthesis, Rittenberg and Bloch (117) have suggested that pyruvate may form acylpyruvic acids and that the decarboxylation to β-keto acids takes place after the condensation

[3] This conclusion was based on experiments with deuterio alanine. It does not take into account the possibility that pyruvate is converted to acetyl by a circuitous route in which deuterium is lost entirely. Recent experiments of Sonne et al. (98) on uric acid synthesis with test substances containing C^{13} similarly indicate that lactate is not converted to acetate.

$$RCH_2COOH + CH_3COCOOH \rightarrow RCH_2COCH_2COCOOH \rightarrow RCH_2COCH_2COOH + CO_2 \quad [9]$$

In this reaction the net effect is the addition of a two carbon unit, without acetic acid appearing as an intermediate. It is also conceivable that an intermediate of the tricarboxylic acid cycle which contains the elements of both acetic and pyruvic acids supplies the C_2 fragments for fatty acid synthesis. For instance, the C_4 dicarboxylic acids which are derived from the condensation product of acetate and oxaloacetate are composed of two C_2 moieties one of which originates from acetate and the other from pyruvate. By passing through a symmetrical stage, randomization would result, and in a utilization of the dicarboxylic acid for fatty acid synthesis an equal chance would exist that the C_2 fragment was contributed by either acetate or pyruvate.

It has been pointed out above that the acetic acid arising in intermediary metabolism appears to be derived primarily from the catabolism of fatty acids and ketogenic amino acids while carbohydrate does not contribute significantly to the acetate pool. It would follow, then, that the acetic acid employed for fatty acid synthesis has its origin in the fatty acids themselves; in other words, the process of acetate utilization merely reflects the continuous molecular regeneration of body fats. Since a large portion of the acetyl fragments arising in fatty acid oxidation is metabolized further and is therefore unavailable for fat resynthesis, carbon has to be provided from another source in order to keep the total quantity of fat constant. The source is evidently carbohydrate. Acetate is therefore not a precursor of fatty acids in the conventional sense and the statement that 25 per cent of the carbon atoms are supplied by acetate merely implies that one out of every four carbon atoms incorporated into a newly synthesized fatty acid molecule arose from the catabolism of the fatty acids themselves.

Cholesterol. In their experiments on the utilization of deuterioacetate by yeast Sonderhoff and Thomas found the deuterium concentration in the unsaponifiable fraction to be twice as high as in the fatty acid fraction and twenty times that of the yeast carbohydrate (127); they therefore suggested that in the synthesis of the yeast steroids acetate had been employed directly. A formation of cholesterol in animal tissues by total synthesis from small units was indicated by the findings of Rittenberg and Schoenheimer (112). In cholesterol newly formed by animals whose body fluids were enriched with heavy water roughly half of the hydrogen atoms were derived from the D_2O of the body fluids. The utilization of various small molecular substances for cholesterol synthesis has been investigated by Bloch and Rittenberg in feeding experiments with labeled compounds (133, 118, 78). The isotope concentration in cholesterol isolated from the tissue lipids was highest when labeled acetic acid was the test substance. Whenever another compound was effective in causing the formation of labeled cholesterol the result could be attributed to intermediary formation of acetate from the test substance (butyrate, ethanol, leucine, valerate, isovalerate). By studying simutaneously cholesterol synthesis and the acetylation of foreign amines it could be shown that the ability of a test substance to cause the formation of

labeled cholesterol was roughly parallel with its ability to furnish labeled acetyl groups. This correlation did not exist in the case of alanine. The significance of this finding has been discussed above. An evaluation of the quantitative rôle of acetic acid as a precursor for the steroid structure led to the conclusion that about half of the hydrogen atoms (and probably half of the carbon atoms) are furnished by acetate. Acetic acid appears to contribute to the formation of the entire steroid molecule. When cholesterol formed biologically from deuterio acetate was degraded, the iso-octyl fragment derived from the cholesterol side chain and the hydrocarbon representing the steroid skeleton contained deuterium in nearly equal concentrations. No information is available as to the intermediate steps of the synthetic process. The higher fatty acids or intermediates of fatty acid metabolism cannot lie on the path of the acetate-sterol conversion, since in feeding experiments with deuterio acetate, newly formed cholesterol contains several times the isotope concentration of the fatty acids; also, in liver slices deuterium is rapidly incorporated into cholesterol when labeled acetate is present, but not into fatty acids (131). The utilization of both carbon atoms of acetic acid is evident from the appearance of both deuterium and C^{13} in cholesterol synthesized from $CD_3C^{13}OOH$ (117, 131).

Since the acetate-cholesterol conversion entails essentially the transformation of—$CD_2 \cdot CO$—into—$CD_2 \cdot CH_2$—two normal hydrogen atoms should be incorporated for every two carbon-bound deuterium atoms and in the utilization of $CD_3C^{13}OOH$ the ratio of D to C^{13} should decline to half its original value. However, the relative amounts of D and C^{13} introduced into cholesterol were roughly the same as in the administered acetate and it would appear therefore that the reaction product contained, perhaps as the result of decarboxylations, more of the methyl carbons than of the carboxyl carbons of acetic acid.[4]

Some indications exist that a reduction may be the initial step in the acetate-cholesterol conversion. McLean and Hoffert (126) observed that sterol synthesis from acetate in yeast was inhibited by sulfite, a reagent which blocks aldehyde groups. In the formation of labeled cholesterol by rat liver slices, labeled ethanol and acetaldehyde were somewhat more effective than acetic acid (97) suggesting that the intermediate is more readily formed from ethanol and acetaldehyde. It is unlikely that the somewhat higher efficiency of acetaldehyde and ethanol under these conditons is due to the fact that they diffuse more readily into the liver cell since the deuterium concentration in cholesterol is independent of the acetate concentration in the buffer medium over a wide range (131).

In the intact animal and in liver slices deuterio alanine fails to give labeled cholesterol (118, 97). Alanine, and presumably pyruvic acid, are therefore not convertible to the C_2 compound which provides carbon for cholesterol synthesis and which is readily formed from acetic acid itself and from the higher

[4] The results obtained with intact animals differ from those obtained in liver slices. The ratio of D to C^{13} in cholesterol formed from acetate in vitro changes from 8 to 2.5. This might be explained by the loss of carbon bound deuterium associated with the interconversion of acetate and acetoacetate which proceeds rapidly in liver slices but apparently not in the intact animal.

fatty acids. This finding supports the previously expressed view that if a C_2 fragment arises in carbohydrate oxidation it is unlikely to be identical with that derived from fatty acids.

The origin of the remaining carbon and hydrogen atoms in cholesterol, which are not accounted for by acetic acid, is unknown. In view of the fact that cholesterol contains a branched side chain and angular methyl groups and that these groupings are otherwise present only in dietary constituents which animals cannot synthesize (valine, leucine, carotene), the possibility has been tested that one of the branched chain amino acids participates in steroid synthesis. The feeding of leucine which contained deuterium at all positions of the carbon chain resulted in the formation of labeled cholesterol while similary labeled valine was ineffective (78). However, leucine is a ketogenic amino acid and in the course of its catabolism yields acetyl groups. Hence, no conclusions may be drawn as to whether leucine specifically supplies the branched side chain and the angular methyl groups of cholesterol or whether the observed effect is ascribable merely to the intermediary formation of acetic acid.

Cholesterol can serve as the parent substance for other compounds which possess the cyclopentano-phenanthrene structure. The in vivo transformation of cholesterol into cholic acid (134) and into pregnanediol (135) has been demonstrated with the aid of deuteriocholesterol and it has been found that a minimum of two-thirds of these steroids can arise by degradation of cholesterol. It is therefore not too likely that animal steroids other than cholesterol can also be formed by total synthesis from small units by a pathway in which cholesterol is not an intermediate.

Protoporphyrin. Hemin isolated from the red cells of animals which receive deuterio acetic acid contains deuterium in significant concentrations (30). In short term feeding experiments the isotope content of hemin will attain only a relatively low level, since the protoporphyrin moiety of hemoglobin is synthesized at a much slower rate than most body constituents (114). When deuterioacetate feeding is extended to a period of several weeks the deuterium concentration in hemin reaches a value corresponding to 50 per cent of the isotope concentration in the acetic acid available in the metabolic pool (136). Therefore at least half of the hydrogen atoms in the porphyrin molecule are supplied by acetic acid. The constituent pyrrol nuclei in the porphyrin molecule contain no hydrogen linked to carbon; carbon bound hydrogen is present in the methin bridges which link the pyrrol rings and in the pyrrol side chains. Hence, experiments with deuterio acetic acid do not reveal whether acetic acid is also a source of carbon atoms for the pyrrol rings. A mechanism of synthesis which leads directly to a pyrrol structure containing the appropriate side chains would seem to be a more likely event than a process in which the substituent side chains are attached to a preformed pyrrol nucleus. No deuterium is introduced into hemin after the feeding of deuterio propionic acid while β-γ-dideuterio butyric acid is about half as effective as deuterio acetic acid (136). The effect shown by butyrate is most likely due to its splitting into acetyl fragments.

The origin of pyrrol nitrogen has been elucidated by the experiments of

Shemin and Rittenberg (114, 137). Hemin isolated from the blood of a human subject who had received glycine labeled by N^{15} contained so high an isotope concentration as to indicate a specific utilization of glycine nitrogen. The specific role of glycine in pyrrol synthesis received further support from the finding that glycine was the most effective source of pyrrol nitrogen among several amino acids tested (114). In particular, proline and glutamic acid, which have been suggested as possible biological precursors of pyrrols because they contain or can readily form heterocycles similar to pyrrols, were much less effective as sources of nitrogen for hemin than glycine. The utilization of the two carbon compounds, acetic acid and glycine, points to a mechanism of porphyrin synthesis in which the ring structure and the substituents are formed simultaneously, possibly by processes analogous to those which yield pyrrols in organic chemistry. Fischer and Fink (138) have obtained evidence that ring closure to a pyrrol occurs when glycine and formylacetone (the aldehyde of acetoacetic acid) react in aqueous solution at 37°. The classical pyrrol synthesis of Knorr (139) involved the condensation of α-amino acetoacetic ester with acetylacetone. Acetic acid can be converted to aceto-acetate in vivo and it is not inconceivable that by an analogous process α amino-acetoacetic acid is formed biologically from acetic acid and glycine. A biological pyrrol synthesis analogous to the Knorr reaction would be attractive because it would afford a product in which the arrangement and the chemical nature of the substituents resembles that in the naturally occurring protoporphyrin.

Uric acid. Recently the utilization of C_2 compounds for the synthesis of still another group of tissue constituents has been reported. Sonne, Buchanan and Delluva (98) have studied the incorporation of C^{13} into uric acid, isolated from pigeon excreta, after the feeding of $CH_3C^{13}OOH$, $CH_3^{13}C^{13}HOHCOOH$, $CH_3CHOHC^{13}OOH$ and $NH_2CH_2C^{13}OOH$ respectively. The carboxyl carbon of acetic acid was located at the two ureido carbons (2 and 8) of the purine structure, while the carboxyl group of glycine supplied carbon for position 4, and CO_2 carbon for position 6. On the basis of the data available so far no plausible mechanism for purine synthesis from CO_2, acetate and glycine can be suggested. It is of interest in this connection that Barker and Beck (140) found ammonia, carbon dioxide, acetic acid and small amounts of glycine to be formed as products of uric acid fermentation by *Clostridium acidi urici*. The findings of Sonne et al. contain important evidence with respect to some other metabolic processes. The appearance of acetate and glycine carbon at different positions of the uric acid eliminates the possibility that in the pigeon glycine and acetic acid are interconvertible. On the basis of nutritional studies, Almquist et al. had contended that the chick is unable to synthesize adequate amounts of glycine and that in deficient diets acetate may serve as substitute for glycine (141). The interconvertibility of the two compounds which is implied by these findings is contradicted by the tracer experiments of Sonne et al. unless metabolic differences exist between the chick and the pigeon.

The isotope data obtained by Sonne et al. furthermore show that the α carbon of lactate is not incorporated into the same positions of uric acid as the carboxyl

carbon of acetate. A breakdown of lactate to acetate did, therefore, not occur to any considerable extent. These results are incompatible with the assumption that the hypothetical two carbon compound arising in pyruvate oxidation is closely related to acetic acid.

A relationship between acetate and lipid metabolism in bacteria has been observed by Guirard, Snell and Williams (142). Acetate stimulates the early growth of lactic acid bacteria, an effect which is also shown by fatty acids and various steroids. These results suggest an important rôle of acetate in the synthesis of lipids by bacterial systems.

As the result of investigations with isotopically labeled test substances there have been developed a number of novel biochemical theories among which Schoenheimer's concept of the dynamic state of the body constituents has been the most outstanding (143). According to this concept, there exists a pool of metabolites which is continuously replenished by the breakdown products of both dietary and tissue constituents and which supplies the metabolites for energy production and for the rebuilding of the structural components of the cell. The concept of the metabolic pool was advanced in order to interpret the data obtained from feeding experiments with isotopic substances and in order to provide a reasonable basis for the rapid interchange of dietary and tissue elements which was indicated by these results. It should be pointed out that the metabolic pool may not possess any physical reality in the sense that there exists in the body at a given moment a large reservoir of metabolites in which molecules derived from dietary and various tissue constituents are distributed in random fashion and are therefore undistinguishable. It is merely permissible to say that if for example 100 molecules of dietary origin enter the tissues during a given period of time and 100 molecules arise from body constituents during the same interval, then the chances that in a subsequent reaction either the exogenous or endogenous molecule is employed will be equal. The metabolic pool should therefore be regarded merely as an expression of this probability.

One of the problems of metabolism which has held the general interest of the biochemist has been the mutual interconversion of fat, carbohydrate and protein. These questions have only partially been answered by balance techniques in which the criterion of a biochemical conversion is the net change in amount of the reaction product which is induced by the precursor. In the light of the concept of the metabolic pool, biochemical interconversions assume a different meaning. Fat, protein and carbohydrate are interconvertible in the sense that they give rise to breakdown products which lose the identity of their origin and can be utilized interchangeably for the resynthesis of either one of the three major tissue constituents.

The fact that acetic acid is employed as a building stone in the synthesis of a variety of compounds of different chemical structure and function, viz. lipids, carbohydrate, porphyrin, uric acid and some of the amino acids, suggests its direct or indirect participation in the biological formation of all compounds which can be synthesized by the animal cell. If this is true then acetic acid cannot be regarded as a specific precursor of a particular tissue constituent, but its utiliza-

tion would reflect merely the fact that the body employs two carbon units as a principal source for synthetic reactions.

The biologically active form of acetic acid. Since in isolated tissues acetic acid shows less reactivity than in the intact animal the possibility has been widely discussed that acetic acid is metabolically converted into a more reactive form. In organic chemistry such activations are achieved by esterification or anhydride formation (acetylchloride, acetic anhydride, ketene, ethylacetate).[5]

Experimental evidence for the biological formation and the participation in a synthetic reaction of an acetic acid derivative rather than of acetic acid itself has been provided in two instances. In bacterial systems the phosphoroclastic splitting of pyruvate to acetylphosphate and formate can be reversed (145). No other function for acetylphosphate has been found so far. A process of intermolecular acetyl transfer in which acetyl groups are shifted from an N-acetyl amino acid to a free amino acid has been indicated by acetylation studies in the intact animal (35).

There are two other well known biochemical processes which are classified as transfer reactions, transfer of methyl groups involving choline, methionine and creatine (146) and the amidine transfer from arginine to glycine in the formation of guanido acetic acid (147, 148). Such reactions cannot be described in terms of ordinary chemical equations involving stable compounds as intermediates. It is conceivable that group or radical transfer may be a more general phenomenon in biological systems and that the answer to the acetyl problem may be sought in this direction.

THE OXIDATIVE METABOLISM OF ACETIC ACID. The ability of animal tissues to utilize acetic acid efficiently is illustrated by the fact that when fed to animals acetic acid does not appear in the urine in measurable amounts. In the experiments of Buchanan, Hastings and Nesbett with acetate containing radioactive carbon in the carboxyl group, over fifty per cent of the acetate absorbed by the rats was totally oxidized within two hours, as determined by the appearance of isotopic carbon in the expired carbon dioxide (120). This rate was comparable to that at which similar quantities of lactic acid were metabolized (149).

Acetate disappears aerobically from a variety of isolated tissues, the rate being most rapid in the cortex of guinea pig-kidney (150). Little evidence exists that the acetate molecule as such can undergo oxidative reactions.[6] Products which could conceivably arise by oxidation of the methyl group of acetic acid such

[5] For a consideration of the "active" form of acetic acid the following organic chemical reactions are of interest. The studies by Bergmann et al. (36) on intermolecular transfer of acetyl groups or radicals between amino compounds have been mentioned above. Kharasch and Gladstone (144) have shown that by conversion into diacetyl peroxide, the methyl groups of acetic acid become sufficiently activated to undergo α-α condensation to succinic acid. The reaction is believed to be due to the dimerization of an intermediate CH_2COOH radical.

[6] The claim of Toenniessen and Brinkmann (151) that formate is a product of acetate metabolism in perfused muscle is based on insufficient experimental evidence, but conforms with recent observations of Buchanan and Sonne (167) on the rôle of acetate and formate respectively as precursors of uric acid.

as glycolic, glyoxylic, oxalic acids or glycine are metabolically inert under conditions which permit of rapid oxidation of acetate (152). Many years ago Thunberg suggested that the oxidation of acetate was initiated by the dehydrogenative coupling of the methyl groups of two molecules of acetate to succinic acid (55).

$$HOOC\cdot CH_2\boxed{H\ +\ H}CH_2COOH \longrightarrow HOOC\cdot CH_2CH_2COOH \quad [10]$$

The same reaction was proposed later by Wieland and Sonderhoff to account for the formation of succinate by yeast in which acetate was the sole nutrient (153). Slade and Werkmann (108) have studied this process with C^{13} containing acetate in *Aerobacter indologenes* and have concluded that the isotope data demonstrate methyl condensation of two moles of acetate. Lipmann has suggested that this condensation may involve acetylphosphate (154). The inhibition by malonate of acetate and acetoacetate disappearance in animal tissues (155, 156) and of acetate in yeast (157) has been cited in support of reaction [10] but the fact that acetate oxidation is sensitive to malonate may merely be taken to suggest that succinic acid lies on the path of acetate metabolism; it does not indicate the mechanism involved. There is increasing evidence that the step which precedes acetate oxidation consists of a condensation with a C_4 dicarboxylic acid. From the resultant C_6 tricarboxylic acid two carbon atoms are removed by oxidation to regenerate the dicarboxylic acid:

$$CH_3COOH + C_4 \rightarrow C_6 \rightarrow 2\,CO_2 + C_4 \quad [11]$$

The experimental support for this scheme coincides in many instances with the evidence which led originally to the formulation of the "citric acid" cycle for the metabolism of pyruvic acid. In fact, in some respects the rôle of the tricarboxylic acids is better understood for the oxidative phase of fatty acid and acetate metabolism than for the oxidation of pyruvate. Acetate and acetoacetate have been shown to form the intermediate polycarboxylic acids of the cycle, but it has remained undetermined whether pyruvate itself enters into combination with the C_4 dicarboxylic acids or whether the condensation is preceded by a degradation of pyruvate to a two carbon unit. The significance of the citric acid cycle in the metabolism of fat and carbohydrate respectively is still a matter of controversy. According to Breusch, carbohydrate is not oxidized to an appreciable extent by way of the citric acid cycle (158), while Weil-Malherbe (159) and Krebs (160) reject the view that the citric acid cycle is involved in the oxidation of acetoacetate.

Acetoacetate, acetate and pyruvate have been shown to be convertible, in the presence of dicarboxylic acids, to citrate and intermediates of the tricarboxylic acid cycle, but whether a single mechanism is invoved in all cases is undecided. However, the citric acid cycle constitutes the most successful attempt to date to describe the oxidative phases of fat, carbohydrate and protein metabolism by a unified scheme which visualizes a junction of metabolic pathways and leads to a common mechanism for the ultimate oxidation of the three major dietary and tissue constituents.

The citric acid cycle in relation to the metabolism of carbohydrate has been discussed in reviews by Werkmann and Wood (161), Krebs (162) and Evans (163). The citric acid cycle as a vehicle for fat oxidation has been considered in detail by Wood (123). The present discussion will be concerned primarily with experimental data which deal with the oxidation of acetic acid.

Since there is little evidence to support the assumption that the carbon chain of acetate as such is susceptible to oxidation the question arises as to the steps which precede the conversion of acetate to carbon dioxide and water. Breusch (158) has suggested that the intermediate β keto acids formed in fatty acid oxidation combine directly with oxalacetate to give a product which contains the two terminal carbon atoms of the fatty acid chain, leaving a fatty acid shortened by two carbon atoms. According to this view, two carbon compounds would not arise as distinct intermediates in fat oxidation. Since acetic acid is a compound of great chemical stability and, compared to other metabolites, relatively unreactive in some biological systems, the view has frequently been expressed (123) (164) that acetic acid itself may never arise but may be the stabilization product of a more reactive C_2 unit. The identity of this intermediate has not been established in animal tissues. Judging from the variety of reactions which it undergoes in the intact animal and in some isolated systems, acetic acid must be readily convertible to the hypothetical reactive C_2 fragment.

$C_2 + C_1$ addition. The reversal of the phosphoroclastic splitting of pyruvic acid to formate and acetylphosphate according to the equation:

$$CH_3COCOOH + H_3PO_4 \rightleftarrows CH_3COOPO_3H_2 + HCOOH(H_2 + CO_2) \quad [12]$$

is well established in bacterial systems (145). Lipmann has demonstrated con vincingly that acetyl phosphate is the product of this bacterial dissimilation (165). The preservation of a large fraction of energy in the carboxylphosphate facilitates the reverse reaction, i.e., the carboxylation of the C_2 compound, a reaction which is energetically improbable with acetate itself. It has not been possible to demonstrate acetylphosphate as a product of pyruvate metabolism in animal tissues nor can the evidence for the formation of acetylphosphate from acetate in animal tissues be considered sufficient. Lipmann and Tuttle (166) have incubated extracts from pigeon liver with acetate and adenosine triphosphate and observed that on addition of hydroxylamine and ferric chloride a chromogen was formed which had the properties of hydroxamic acids obtainable from acetylphosphate or aldehydes, but more direct proof is desirable to establish the identity of the acetyl compound formed under these conditions. The reversal of reaction [12] is a formylation rather than a carboxylation and apparently CO_2 can be utilized only in organisms which hydrogenate carbon dioxide to formate. In considering the likelihood that reaction [12] occurs in animal tissues, it should be noted that so far no function has been assigned to formic acid in animal metabolism.[7] Experiments of Lorber et al. (121) in which the isotope distribu-

[7] This statement may have to be revised in view of the recent finding by Buchanan and Sonne (167) that formate is utilized in uric acid synthesis by pigeons. The authors suggest that formic acid may arise by oxidation of acetate.

tion in glucose was studied after feeding of labelled acetic acid contain no evidence for $C_2 + C_1$ addition. Carboxylation of acetic acid which contained C^{13} in the carboxyl group would yield pyruvate with the labeled carbon at the α carbon atom. If such pyruvate were further carboxylated to oxaloacetate and reformed after passing through the stage of a symmetrical dicarboxylic acid, the isotopic carbon would become equalized between the α and β positions of pyruvate. Glucose formed from the singly labeled pyruvate should contain isotope at positions 2 and 5 and, if formed from doubly labeled pyruvate, in positions 1 and 6 as well. Actually a significant excess of C^{13} was encountered only at positions 3 and 4. Thus a carboxylation of acetate to pyruvate cannot be responsible for the incorporation of acetate carbon into the glucose units of glycogen. Data pertaining to the formation of glycogen or intermediates of the citric acid cycle in the presence of labeled CO_2 do not permit any deductions as to the occurrence of acetate carboxylation because any labeled carbon entering by the latter reaction becomes indistinguishable from the carbon which is incorporated into the C_4 dicarboxylic acids by the primary carboxylation of pyruvate to oxalacetate.

$C_2 + C_2$ addition. There exist several observations which are consistent with Thunberg's suggestion (55) that succinate might arise by the linking of the methyl groups of two moles of acetate. The conversion of acetate to succinate by yeast with acetate as the sole nutrient was attributed by Wieland and Sonderhoff to the dehydrogenative coupling of acetate (153). In an analogous experiment with deuterio acetate Sonderhoff and Thomas (127) found high concentrations of deuterium in succinate and citrate and interpreted their results as evidence for reaction [10]. The proof for this contention would be conclusive only if the succinate contained close to four atoms of deuterium instead of the two atoms actually found. More than one pathway can account for the observed level of isotope in succinic acid. The same qualification applies to the findings on succinate formation from labeled carbon containing acetate by *Aerobacter indologenes* (108). The possibility that in both cases succinate arose by way of the tricarboxylic acid cycle cannot be excluded. At the present moment no experiments are available which establish the occurrence of reaction [10] in the conversion of acetate to succinate. Whenever succinate formation from acetate is demonstrable, succinate is not the sole product but α-ketoglutarate and tricarboxylic acids also accumulate. If acetate were metabolized according to equation [10] a mechanism should also be available for the subsequent conversion of succinate to ketoglutarate and citrate. The discovery of the carboxylation of α-ketoglutarate to oxalosuccinate by Ochoa (168) provides a pathway to the tricarboxylic acids, but the reversibility of the oxidative decarboxylation of ketoglutarate to succinate remains to be demonstrated.

$C_2 + C_3$ addition. The possibilities have been considered that acetic and pyruvic acids combine to form the five carbon acids α-ketoglutaric and aceto-pyruvic acids. As a result of the current prominence of the tricarboxylic acid cycle these two processes, which would lead directly to five carbon acids, have received diminishing attention. Krebs and Johnson (169) have suggested the formation of aceto-pyruvate according to the equation

$$CH_3CO|\underline{OH + H}|CH_2COCOOH \rightarrow CH_3COCH_2COCOOH$$
$$CH_3COCH_2COCOOH + \tfrac{1}{2}O_2 \rightarrow CH_3COCH_2COOH + CO_2 + H_2O \quad [13]$$

as a conceivable intermediate for the conversion of pyruvate to acetoacetate in liver. The utilization of acetopyruvate by intact animals and in isolated tissues has been studied by Lehninger (170). The compound was found to be ketogenic. Since acetoacetate can be metabolized by way of the tricarboxylic acid cycle the reaction offers a chemically plausible scheme for the conversion of pyruvate to the oxidation level of acetate.

Acetopyruvate cannot be an obligatory intermediate in ketone body formation from fatty acids. Weinhouse et al. (76), in a study of the oxidation of labeled octanoic acid by liver, found an isotope distribution in acetoacetate which eliminated the participation of pyruvate as an essential component. It is however conceivable that ketone bodies can be formed by two independent synthetic processes, by condensation of two molecules of acetate and by combination of acetate with pyruvate.

α-Ketoglutarate is a key substance in the tricarboxylic acid cycle and a demonstrable intermediate whenever fatty acids or pyruvate are oxidized. A process leading to α-ketoglutarate by "end to end" condensation of acetate and pyruvate has been suggested by Krebs (171):

$$HOOCCH_2|\underline{H + H}|CH_2COCOOH \rightarrow HOOCCH_2CH_2COCOOH \quad [14]$$

Weinhouse et al. have observed the formation of labeled α-ketoglutarate from isotopic acetate and normal oxaloacetate in rat kidney slices or mince, but found no evidence for the simultaneous formation of marked tricarboxylic acids (172). In their experiments non-isotopic citrate and cis-aconitate were added as carriers to facilitate the recovery of citrate at the end of the incubation period. The failure of acetate carbon to become incorporated into citrate in spite of the presence of aconitase, led the authors to the conclusion that citric acid as such was not in the pathway of acetate oxidation by kidney. These results would seem to suggest that α-ketoglutarate arose from acetate and pyruvate according to equation [14] without intervention of the tricarboxylic acids. However in this case the data exclude rigidly only the formation of citric acid itself. It does not appear to be feasible to isolate isocitric acid or cis-aconitate from such experiments. That the latter two tricarboxylic acids may have been intermediates becomes clear from kinetic data of Martius (173) and Krebs (174). These authors have shown that in systems containing the tricarboxylic acids and aconitase the conversion of cis-aconitate to isocitrate proceeds much more rapidly than to citrate. As the conversion of isocitrate to α-ketoglutarate is also a rapid reaction it is possible that in the experiments of Weinhouse et al. cis-aconitate or isocitrate were the initial reaction products but that the breakdown to α-ketoglutarate proceeded at a much faster rate than isomerization to citrate.

In a more recent investigation (175) the same authors report that after incubation of kidney with isotopic acetoacetate and normal oxaloacetate, without addition of non-isotopic carrier, citrate could be isolated directly and shown to

contain an isotope level which indicated that about two-thirds of the citrate arose from the isotopic acetoacetate. These recent experiments differed from earlier ones only in that acetoacetate instead of acetate was the substrate. In view of the fact that acetate is readily convertible to acetoacetate under these conditions it is not apparent why labeled citrate should not have formed in both instances.

In an analysis of the available data pertaining to the citric acid cycle, Wood (123) points out that α-ketoglutarate formation according to reaction [14] as opposed to its formation from tricarboxylic acids cannot be ruled out at the present. If reaction [14] were responsible for α-ketoglutarate formation rather than the oxidative decarboxylation of isocitric acid, then the formation of tricarboxylic acids would have to be ascribed to the carboxylation of α-ketoglutarate to oxalosuccinate and reduction of the latter to isocitrate. Experiments with labeled test substances and studies of isotope distribution in the α-ketoglutarate formed from either labeled acetate or acetoacetate have been incapable of providing an answer as to which reaction occurs. Labeled carbon will occupy the same positions whether α-ketoglutarate is formed by oxidative decarboxylation of isocitrate, directly from acetate and pyruvate or by carboxylation of succinate formed according to equation [10]. The condensation of oxaloacetate and acetate to a tricarboxylic acid does not as such involve an oxidation but can be formulated either as an addition if isocitrate is the primary product, or as a condensation with elimination of water if cis-aconitate is initially formed. It differs in this respect from reactions [10] and [14] in which the coupling process is oxidative. Experimentally it has been found that the formation of the intermediates of the citric acid cycle requires aerobiosis (176) unless a hydrogen acceptor such as oxaloacetate is present in excess (177). This may indicate the occurrence of oxidative condensations or, as is generally believed, oxygen may be necessary to effect the transformation of the reactants into activated forms or derivatives.

There is only scant information on the mechanisms by which carbon-carbon bonds are established biologically. It may be pointed out that in the few processes in which the chemical identity of the reacting molecules seems well established, e.g., in the formation of hexose from triose, or of acetoacetate from acetate, or in the carboxylation of pyruvate to oxaloacetate, at least one of the reacting groups is a carbonyl group. No instances are known of a formation of carbon to carbon bonds by elimination of hydrogen between methyl- or methylene groups.

$C_2 + C_4$ condensation. At present the ultimate oxidation of the split products of fatty acids can be explained most satisfactorily by assuming the intermediate formation of a tricarboxylic acid according to equation [11] and the subsequent breakdown of the condensation product by way of α-ketoglutaric acid and the C_4 dicarboxylic acids. Reaction [11] was originally proposed by Virtanen (178) to account for citrate formation from acetate in micro-organisms. The following evidence has been secured in support of the suggested rôle of the citric acid cycle for fat and acetate oxidation in various biological systems. Wieland and Sonderhoff (153) found succinate and citrate to accumulate as products of acetate

oxidation by yeast when acetate was the sole nutrient. Under these conditions, the yield of citrate is increased by the addition of oxaloacetate (179). Succinate and citrate formed from deuterioacetate contained 50 per cent and 46 per cent respectively of the deuterium concentration of acetate in stably bound form (127). Lynen has contributed important evidence to show that the formation of succinate from acetate in yeast involves the tricarboxylic acid cycle (157). He demonstrated that the malonate inhibition of acetate disappearance could be overcome by fumarate or oxalacetate and cites this finding as evidence for a requisite participation of the C_4 dicarboxylic acids in acetate oxidation. The observation that the rapid disappearance of acetic acid in yeast is preceded by an induction period was interpreted to show that a second component must accumulate before acetate can be metabolized at the maximum rate. A stimulatory effect of citrate, α-ketoglutarate and the C_4 dicarboxylic acids has been observed for the aerobic disappearance of acetoacetate in kidney homogenates by Buchanan et al. (180). Wieland and Rosenthal (176) obtained optimum yields of citrate in kidney brei when acetoacetate was oxidized in the presence of oxaloacetate. The identity of the reaction product with citric acid was established by direct isolation. Hunter and Leloir (177) have shown that two molecules of extra citric acid arise from each mole of acetoacetate. They have made the observation that in their experimental system, which consisted of the insoluble particles of kidney cortex homogenate, the simultaneous oxidation of α-ketoglutarate was essential for citrate formation. This oxidative process is believed to furnish the necessary energy for the conversion of acetoacetate to the intermediate which condenses with oxaloacetate. Lehninger (89) has found that the oxidation of octanoic acid in a preparation of washed rat liver cells yields acetoacetate quantitatively, but that in the presence of fumarate, citrate and α-ketoglutarate accumulate in significant quantities at the expense of acetoacetate. According to Breusch (181) an enzyme can be extracted from kidney, brain and skeletal muscle which catalyzes the anaerobic condensation of β-keto acids with oxaloacetate to citrate.

In addition to the balance studies a number of experiments have been carried out with isotopically labeled substances. Buchanan et al. (180) incubated kidney homogenates with $CH_3C^{13}OOH$ or acetoacetate marked by heavy carbon at both carboxyl and carbonyl positions. Non-isotopic ketoglutaric acid, fumarate or succinate were added as carriers; when re-isolated after incubation these acids were shown to contain C^{13} in the carboxyl positions. It was calculated that the γ-carboxyl of α-ketoglutarate contained almost ten times as much C^{13} as the carboxyl carbon adjacent to the carbonyl group. Essentially the same findings were obtained by Weinhouse et al. (172). The formation of labeled citrate from acetoacetate recently demonstrated by the same authors (175) has been mentioned above.

Intermediates of the citric acid cycle can be obtained from isolated tissues on incubation with a variety of substrates but since they are short-lived intermediates *in vivo* they are not obtainable as such in sufficient amounts from intact animals. It has therefore not been possible to secure direct evidence as to

whether the tricarboxylic acid cycle is concerned in fatty acid oxidation also in the intact animal. Feeding experiments have been carried out by Buchanan, Hastings and Nesbett (120) with acetic, propionic and butyric acids, all of which contained C^{11} in the carboxyl group. Glycogen isolated from such animals contained radioactive carbon in concentrations which in the case of acetate feeding were attributable to incorporation of CO_2, but were significantly higher with radioactive propionate and butyrate. Evidently the latter two acids had furnished carbon for glycogen synthesis by conversion to intermediates of carbohydrate metabolism. With the aid of degradation procedures which allow isotope determinations at individual carbon atoms of the glucose molecule (115), Wood and his collaborators have been able to show that the carboxyl carbons of acetic acid appear at positions 3 and 4 (121) and the methyl carbons at positions 1, 2, 5 and 6 of the glucose molecule (182). Thus, both carbon atoms of acetic acid are involved in glycogen formation. Butryric acid $CH_3C^{13}H_2CH_2C^{13}OOH$ afforded the same isotope distribution in glycogen as $CH_3C^{13}OOH$ (182). From these findings the conclusion can be drawn that there exists a pathway for the conversion of the two carbon split products of fatty acids into the glucose units of glycogen. In this process each carbon atom of glucose can originate from fatty acid carbon.

Glutamic and aspartic acids are believed to be in biological equilibrium with α-ketoglutarate and oxaloacetic acid. Any incorporation of marked fatty acid carbon into intermediates of the tricarboxylic acid cycle should therefore become manifest by the formation of labeled aspartic and glutamic acids. The two amino acids have been isolated by Rittenberg and Bloch from the tissue proteins of mice which had received $CH_3C^{13}OOH$, and were shown to contain significant concentrations of C^{13} (183).

While the evidence which has accumulated in favor of the tricarboxylic acid cycle is impressive it should be emphasized that this scheme does not offer a unique explanation for the formation of tri- and dicarboxylic acids as products of fatty acid oxidation. The following points of the cycle need further clarification: 1, the identity of the intermediate of fatty acid and pyruvate metabolism which condenses with oxaloacetate; 2, the nature of the initial condensation product; 3, the existence of one or more mechanisms for the oxidation of acetate, acetoacetate and pyruvate respectively.

Although the stimulatory effect of the C_4 dicarboxylic acids on citrate formation is exhibited not only by oxaloacetate but also by succinate, malate and fumarate, it is reasonable to asume that the effect is attributable in all cases to oxaloacetate. The nature of the acetyl compound which participates in the condensation process is less evident. In some systems intermediates of the cycle are formed from acetate as well as from acetoacetate, while in other situations the two substances cannot be employed interchangeably. In balance studies with kidney preparations (176, 177) and in the experiments of Lehninger (66) with a heart muscle preparation only acetoacetate was an effective precursor of citrate. The failure of acetate to react in kidney preparations is unexpected because acetate is readily converted to acetoacetate in this organ. On the other hand,

Buchanan et al. (180) found that labeled acetate was as effective as acetoacetate in forming labeled α-ketoglutarate in kidney homogenate. It should be reemphasized in this connection that with the aid of the tracer technique a conversion can be demonstrated even if there is no net change in the concentration of the reaction product, e.g., citrate or α-ketoglutarate. This is important in systems in which the product is metabolized further at a rapid rate.

In general, acetoacetate appears to be the more reactive molecule with regard to the condensation reaction. On the basis of their data Wieland and Rosenthal (176) postulated that acetoacetate itself condensed to yield the hypothetical procitric acids, citroylacetic or acetylcitric acids. Weinhouse et al. were able to isolate C^{13} containing citrate from kidney homogenate when acetoacetate was the substrate (175), while in earlier experiments under similar conditions, labeled acetate yielded isotopic α-ketoglutarate but not citrate (172). The preferential utilization of acetoacetate might be taken to indicate that acetate enters the tricarboxylic acid cycle by way of acetoacetate. However, several situations exist for which this explanation does not hold. Citrate is a product of acetate oxidation in yeast, but as far as is known, the conversion of acetate to acetoacetate does not occur in this organism. Medes et al. have shown that acetate undergoes complete oxidation in heart muscle but that acetoacetate is not an intermediate (86). In kidney also a portion of acetate was oxidized by a route which did not pass through the acetoacetate stage. The data allow for two possible explanations: 1, acetate and acetoacetate enter the tricarboxylic acid cycle by a different process, and 2, a reactive acetyl compound is formed from both acetate and acetoacetate before condensation occurs. In the utilization of acetoacetate a preliminary splitting to acetic acid can be excluded as an intermediate step. Buchanan et al. (184) added normal acetate and isotopic acetoacetate to an extract of rabbit kidney and found that the acetic acid recovered from the mixture after incubation contained appreciably less isotope than α-ketoglutarate or succinate. From energy considerations it is to be expected that an energy rich acetyl can arise far more readily from acetoacetate than from acetate. The inability of acetate to substitute for acetoacetate may be ascribed to an impairment of the enzyme system responsible for the conversion of acetate to the metabolically active form. Although acetoacetate is an efficient source of citrate in various tissue systems, it is highly doubtful that it is an obligatory intermediate in tricarboxylic acid formation from fatty acids. In the experiments of Lehninger with washed liver cells, acetoacetate, unlike octanoic acid, was incapable of yielding citrate (89). The preparation apparently had lost the property of converting acetoacetate into the active two carbon fragment. The reasons which led to the belief that in the intact animal acetoacetate is not an intermediate of fatty acid oxidation have been discussed above.

Since the experimental conditions employed by various investigators for the study of the same problem vary widely, particularly with respect to methods of tissue preparations, it is not surprising that such experiments have often yielded contradictory results and that it is not possible to fit all available data into a single scheme. It is evident that with increasing impairment of the cellular

organization the enzymatic complement of the biological system becomes increasingly deficient and hence the probability will diminish that a complex series of reactions can take place. Processes which have high energy demands and must be coupled with energy yielding reactions such as the conversion of acetate to an active acetyl are less likely to occur in homogenates or tissue extracts than in systems in which the cell structure is preserved.

Lipmann has shown that by the addition of adenosine triphosphate acetate can be activated to acetylate sulfanilamide anaerobically in pigeon liver extracts (2). The expectations, based on this finding, that the active C_2 compound which is involved in the tricarboxylic acid cycle may be identical with acetyl phosphate, have not been fulfilled. Under conditions in which citrate formation can be demonstrated to occur with acetoacetate but not from acetate, acetylphosphate is also ineffective (184, 177, 89). In order to determine the identity of the active two carbon compound, Lehninger (89) has also tested the effect of glycolate, glyoxalate, oxalate, acetaldehyde, ethanol, acetamide and glycine. None of the compounds yielded extra citrate under conditions which afforded citrate from octanoic acid.

Since *cis-aconitate, isocitrate and citrate are biologically interconvertible*, the identity of the tricarboxylic acid initially formed in the citric acid cycle has remained in doubt. In the presence of aconitase the system will contain all three acids irrespective of the nature of the primary condensation product (173, 185). The cycle as originally proposed by Krebs and Johnson (185), in which citric acid was the initial condensation product, had to be modified when it was found by Evans and Slotin (186) and by Wood et al. (187) that the α-ketoglutarate formed in pigeon liver homogenate from labelled CO_2 and pyruvate contained labeled carbon only in the carboxyl group adjacent to the carbonyl group, and that when malonate was present, the succinate formed contained little C^{13} (186). The tricarboxylic acids could therefore not have originated from a symmetrical citric acid. An unsymmetrical molecule must also have been the precursor of the dicarboxylic acids which Buchanan et al. (180) and Weinhouse et al. (175) obtained as the reaction product of labeled acetoacetate or acetate, and oxaloacetate. In the α ketoglutarate obtained from the latter experiments only one-sixth and one-fourth respectively of the total isotope was present in the carboxyl adjacent to the keto group; it should have contained half of the total, if a symmetrical tricarboxylic acid had been the precursor. Likewise, in glutamic acid isolated from the proteins of animals which had received carboxyl labeled acetic acid the heavy carbon was not evenly distributed between the two carboxyl groups (183). The C^{13} content of the carboxyl adjacent to the amino-group accounted for only one-fourth of the total, whereas it should have contained half of the total if the α-ketoglutarate arose from a symmetrical precursor. Thus, in the intact animal also, the α-ketoglutarate formed in the process of acetate oxidation is derived from an intermediate in which the orientation of the labeled carbon is preserved. The much higher isotope content of glutamic than of aspartic acid isolated from this experiment suggests that only a small fraction of the isotope incorporated into the glutamic acid was contributed by CO_2 assimila-

tion. The carbon chain of aspartic acid is believed to be furnished by oxalacetate. If isotope had entered exclusively in the form of CO_2 by the carboxylation of pyruvate to oxalacetate, then aspartate should contain not less but as much or more heavy carbon than the α-ketoglutaric acid which is subsequently formed from oxalacetate by way of a tricarboxylic acid.

The results obtained with labeled substances eliminate a symmetrical molecule as the initial condensation product, but do not indicate which tricarboxylic acid is involved. In addition to isocitrate and cis-aconitate, the two acids which according to present views are the most likely intermediates, oxalosuccinate has become a possible intermediate since Ochoa (168) has demonstrated reversibility of the isocitrate-oxalosuccinate conversion. Citric acid has a symmetrical carbon chain but the possibility still exists that a citric acid which is substituted at one of the two primary carboxyl groups lies on the main path of the cycle. The suggestion of Wieland and Rosenthal (176) that the citric acid formed in kidney from acetoacetate might have arisen secondarily from a procitric acid containing eight carbon atoms has been mentioned. Lynen (188) finds the formation of citrate from acetate and oxalacetate by yeast to be promoted by the simultaneous oxidation of succinaldehyde and believes that the function of this reaction is to convert either one of the components into a more reactive derivative. If the condensation takes place with an acetyl compound CH_3 COR, a substituted unsymmetrical citrate molecule may result, which could be converted to cis-aconitate without elimination of the substituent R. Lynen has pointed out that the data of Sonderhoff and Thomas (127) on the formation of labeled citrate and succinate in yeast, also rule out a symmetrical citric acid. Succinate formed from trideuterioacetate by way of a symmetrical tricarboxylic acid could not have contained more than one atom of deuterium, but was actually found to contain two atoms. The suggestion that an unsymmetrically substituted citric acid derivative is initially formed would also account for the unequal distribution of isotopic carbon in α-ketoglutarate formed from labeled acetate or acetoacetate. The question concerning the identity of the initial reaction product might be clarified materially if it were possible to establish whether citrate formation occurs only in systems which contain the enzyme aconitase. According to Krebs (162) this enzyme is absent from yeast. In this case the citric acid arising in the course of acetate oxidation by this organism could not have been formed secondarily from isocitrate or cis-aconitate but must have been the primary condensation product.

A discussion of the rôle of the citric acid cycle in carbohydrate metabolism is beyond the scope of this review, but it may be pointed out that no evidence exists to support the contention that the tricarboxylic acid formed in the oxidative metabolism of pyruvate is identical with that arising from fatty acid oxidation. The finding of Martius (189) that oxalocitramalate, the condensation product of pyruvate and oxalacetate, is not attacked by kidney tissue is hardly sufficient evidence to rule out the occurrence of intermediate C_7 tricarboxylic acids. As the experience with yeast has shown (190, 157), the lack of metabolic activity of a substance added to a biological system may be apparent only and

may be caused by inability of the substance to penetrate cell walls. This is particularly true for polycarboxylic acids. Furthermore, there are other C_7 acids which could conceivably arise from pyruvate and oxaloacetate and which have not been tested biologically (formylcitric, oxalocitraconic acids).

The conclusion that the tricarboxylic acid cycle provides an important mechanism for the complete oxidation of fatty acids in the intact animal is based on evidence of indirect nature. The appearance of isotopic carbon in the glycogen of animals fed labeled fatty acids and in the dicarboxylic amino acids after feeding of acetate is not by itself indicative of the mechanism responsible, but is most reasonably ascribed to pathways which lead to the intermediary formation of tricarboxylic acids. It is not possible at present to determine what portion of the fatty acids is oxidized in the intact animal by this route. From the high isotope level in glycogen in the experiments of Buchanan et al. (120) and of Lorber et al. (121) it is clear however that a significant fraction of acetate and butyrate respectively was metabolized by way of the tricarboxylic acid cycle. In these experiments with starved animals, ordinary glucose was given simultaneously with the labeled fatty acids in order to insure glycogen deposition. The labeled glycogen arising from the labeled fatty acids was therefore diluted by glycogen formed directly from dietary glucose.

The finding of Wood et al. (182) that the same isotope distribution in glucose obtains with butyrate labeled at carbon atoms 1 and 3 as with carboxyl labeled acetate demonstrates the breakdown of butyrate to two acetyl groups and furnishes additional evidence for the view that fatty acids are degraded by successive removal of two carbon units. These data are not in agreement with the suggestion of Blixenkrone-Moeller (191) that butyrate might be oxidized to succinate directly by ω oxidation.

The results obtained during the past few years have not only provided the basis for a reasonable scheme of the total oxidation of fatty acids but have also clarified the controversial relationship of fat and carbohydrate metabolism. From the findings discussed above it may be concluded that fatty acids are convertible to carbohydrate in a restricted sense. The split products of fatty acid oxidation contribute to carbohydrate synthesis but the fatty acids by themselves are incapable of causing an increased formation of glucose or glycogen because the mechanism by which the fatty acid carbons find their way into glucose requires the participation of carbohydrate itself. This explains why in the starved or diabetic animal fatty acids and acetate do not show a glycogenic effect.

REFERENCES

(1) Klein, J. R. and J. S. Harris. J. Biol. Chem. 124: 613, 1938.
(2) Lipmann, F. J. Biol. Chem. 160: 173, 1945.
(3) Loeb, A. Biochem. Ztschr. 47: 118, 1912.
(4) Monguio, J. Klin. Wchnschr. 31: 1116, 1934.
(5) MacKay, E. M., R. H. Barnes, H. O. Carne and A. N. Wick. J. Biol. Chem. 135: 157, 1940.
(6) Bray, H. G., F. C. Neale and W. V. Thorpe. Biochem. J. 40: 406, 1946.
(7) Litchfield, J. T. J. Pharmacol. 67: 212, 1939.

(8) Knoop, F. Ztschr. physiol. Chem. 67: 489, 1910.
(9) Marshall, E. K., Jr., K. Emerson, Jr. and W. C. Cutting. J. A. M. A. 110: 252, 1938.
(10) Stekol, J. A. Ann. Rev. Biochem. 10: 265, 1941.
(11) Handler, P. and W. A. Perlzweig. Ann. Rev. Biochem. 14: 617, 1945.
(12) Green, D. E., D. H. Moore, V. Nocito and S. Ratner. J. Biol. Chem. 156: 383, 1944.
(13) Cohn, R. Ztschr. physiol. Chem. 17: 284, 1893.
(14) Neubauer, O. and O. Warburg. Ztschr. physiol. Chem. 70: 1, 1910.
(15) Knoop, F. and J. G. Blanco. Ztschr. physiol. Chem. 146: 267, 1925.
(16) du Vigneaud, V. and C. E. Meyer. J. Biol. Chem. 98: 295, 1932.
(17) du Vigneaud, V. and O. J. Irish. J. Biol. Chem. 122: 349, 1937.
(18) Erlenmeyer, E. and J. Kunlin. Annalen 307: 146, 1899.
(19) de Jong, A. W. K. Rec. trav. chim. Pays-bas 19: 259, 1900.
(20) du Vigneaud, V., M. Cohn, G. B. Brown, O. J. Irish, R. Schoenheimer and D. Rittenberg. J. Biol. Chem. 131: 273, 1939.
(21) Bernhard, K. Ztschr. physiol. Chem. 267: 91, 1940.
(22) du Vigneaud, V., J. L. Wood and O. J. Irish. J. Biol. Chem. 129: 171, 1939.
(23) du Vigneaud, V., J. L. Wood and F. Binkley. J. Biol. Chem. 138: 369, 1941.
(24) Doisy, E. A., Jr. and W. W. Westerfeld. J. Biol. Chem. 149: 229, 1943.
(25) Martin, G. J. and E. H. Rennebaum. J. Biol. Chem. 151: 417, 1943.
(26) James, G. V. Biochem. J. 33: 1688, 1939.
(27) Bernhard, K. Ztschr. physiol. Chem. 267: 99, 1940.
(28) Bernhard, K. Ztschr. physiol. Chem. 273: 31, 1941.
(29) Bloch, K. and D. Rittenberg. J. Biol. Chem. 155: 243, 1944.
(30) Bloch, K. and D. Rittenberg. J. Biol. Chem. 159: 45, 1945.
(31) Bloch, K. and E. Borek. J. Biol. Chem. 164: 483, 1946.
(32) Krebs, H. A. Ann. Rev. Biochem. 5: 254, 1936.
(33) Annau, E. Ztschr. physiol. Chem. 224: 141, 1934.
(34) Edson, N. L. Biochem. J. 29: 2082, 1935.
(35) Bloch, K. and D. Rittenberg. Fed. Proc. 5: 122, 1946.
(36) Bergmann, M., V. du Vigneaud and L. Zervas. Ber. Chem. Ges. 62: 1909, 1929.
(37) Bergmann, M. and L. Zervas. Ztschr. physiol. Chem. 175: 145, 1928.
(38) Borsook, H. and H. M. Huffman. In C. L. A. Schmidt. The chemistry of the amino acids and proteins. p. 865, 2nd ed., Springfield, 1944.
(39) Zief, M. and J. T. Edsall. J. Am. Chem. Soc. 59: 2247, 1937.
(40) Neuberger, A. and F. Sanger. Biochem. J. 37: 515, 1943.
(41) Berg, C. P., W. C. Rose and C. S. Marvel. J. Biol. Chem. 85: 207, 1932.
(42) du Vigneaud, V., R. R. Sealock and C. Van Etten. J. Biol. Chem. 98: 565, 1932.
(43) Carter, H. E., P. Handler, F. Binkley, H. Fishbeck, W. Risser and J. Weisinger. Proc. Soc. Biol. Chem. 32: xx, 1938.
(44) Neuberger, A. and T. A. Webster. Biochem. J. 40: 576, 1946.
(45) Abbot, L. D., Jr. J. Biol. Chem. 145: 241, 1942.
(46) Jackson, R. W. and J. P. Chandler. Ann. Rev. Biochem. 8: 261, 1939.
(47) Weismann, N. and R. Schoenheimer. J. Biol. Chem. 140: 779, 1941.
(48) Berg, C. P. J. Nutrition 12: 671, 1936.
(49) Berg, C. P. J. Biol. Chem. 104: 373, 1934.
(50) du Vigneaud, V., H. S. Loring and H. A. Craft. J. Biol. Chem. 107: 519, 1934.
(51) Berg, C. P. and M. Potgieter. J. Biol. Chem. 94: 661, 1932.
(52) Rose, W. C. Science 86: 298, 1937.
(53) Knoop, F. Beitr. chem. Physiol. u. Path. 6: 150, 1905.
(54) Dakin, H. D. J. Biol. Chem. 6: 221, 1909.
(55) Thunberg, T. Skand. Arch. Physiol. 40: 1, 1920.
(56) McClendon, J. F. J. Biol. Chem. 154: 357, 1944.
(57) Cook, R. P. and K. Harrison. Biochem. J. 30: 1640, 1936.

(58) Stadie, W. C., J. A. Zapp and F. D. W. Lukens. J. Biol. Chem. 137: 75, 1941.
(59) Lehninger, A. L. J. Biol. Chem. 143: 147, 1942.
(60) Stadie, W. C. Physiol. Rev. 25: 395, 1945.
(61) Rittenberg, D. and G. L. Foster. J. Biol. Chem. 133: 737, 1940.
(62) Lorber, V., N. Lifson, H. G. Wood and J. Barcroft. J. Physiol. 145: 557, 1946.
(63) Himwich, H. F., W. Goldfarb, N. Rakerten, L. H. Nahum and D. Du Bois. Am. J. Physiol. 110: 352, 1934.
(64) Drury, D. R. and P. D. McMaster. J. exper. Med. 49: 765, 1929.
(65) Stadie, W. C. J. Clin. Investigation 19: 843, 1940.
(66) Lehninger, A. L. J. Biol. Chem. 165: 131, 1946.
(67) Bernhard, K. and F. Bullet. Helv. chim. acta 26: 1185, 1943.
(68) Leloir, L. F. and J. M. Muñoz. Biochem. J. 33: 734, 1939.
(69) Butts, J. S., C. H. Cutler, L. F. Hallmann and H. J. Deuel, Jr. J. Biol. Chem. 109: 597, 1935.
(70) Deuel, H. J., Jr., L. F. Hallmann and J. S. Butts. J. Biol. Chem. 116: 621, 1937.
(71) Quastel, J. H. and A. H. M. Wheatley. Biochem. J. 27: 1753, 1933.
(72) Jowett, M. and J. H. Quastel. Biochem. J. 29: 2159, 1935.
(73) Cohen, P. P. J. Biol. Chem. 119: 333, 1937.
(74) Lehninger, A. L. J. Biol. Chem. 157: 363, 1945.
(75) Schoenheimer, R. and D. Rittenberg. J. Biol. Chem. 120: 155, 1937.
(76) Weinhouse, S., G. Medes and N. F. Floyd. J. Biol. Chem. 155: 143, 1944.
(77) MacKay, E. M., R. H. Barnes, H. O. Carne and A. N. Wick. J. Biol. Chem. 136: 503, 1940.
(78) Bloch, K. J. Biol. Chem. 155: 255, 1944.
(79) Butts, J. S., H. Blunden and M. S. Dunn. J. Biol. Chem. 120: 284, 1937.
(80) Rose, W. C., T. E. Johnson and W. T. Haines. J. Biol. Chem. 145: 679, 1942.
(81) Friedmann, F. Biochem. Ztschr. 55: 436, 1913.
(82) Swendseid, M. E., R. H. Barnes, A. Hemingway and A. O. Nier. J. Biol. Chem. 142: 47, 1942.
(83) Medes, G., S. Weinhouse and N. F. Floyd. J. Biol. Chem. 157: 751, 1945.
(84) Weinhouse, S., G. Medes and N. F. Floyd. J. Biol. Chem. 158: 411, 1945.
(85) Medes, G., S. Weinhouse and N. F. Floyd. J. Biol. Chem. 157: 35, 1944.
(86) Medes, G., N. F. Floyd and S. Weinhouse. J. Biol. Chem. 162: 1, 1945.
(87) Morehouse, M. G. J. Biol. Chem. 129: 769, 1939.
(88) Morehouse, M. G. and H. J. Deuel, Jr. Proc. Soc. Exper. Biol. 45: 96, 1941.
(89) Lehninger, A. L. J. Biol. Chem. 161: 413, 1945.
(90) Krebs, H. A. and W. A. Johnson. Biochem. J. 31: 645, 1937.
(91) Weil-Malherbe, H. Biochem. J. 31: 2202, 1937.
(92) Long, C. Biochem. J. 32: 1711, 1938.
(93) Lipmann, F. Skand. Arch. Physiol. 76: 255, 1937.
(94) Lipmann, F. Cold Spring Harbor Symposia Quant. Biol. 7: 248, 1939.
(95) Barron, E. S. G., G. R. Bartlett and G. Kalnitsky. Fed. Proc. 5: 120, 1946.
(96) Evans, E. A., Jr. Biochem. J. 34: 829, 1940.
(97) Rittenberg, D., E. Borek and K. Bloch. Fed. Proc. 5: 151, 1946.
(98) Sonne, J. C., J. M. Buchanan and A. M. Delluva. J. Biol. Chem. 166: 395, 1946.
(99) Silvermann, M. and C. H. Werkmann. J. Biol. Chem. 138: 35, 1940.
(100) Green, D. E., W. W. Westerfeld, B. Vennesland and W. E. Knox. J. Biol. Chem. 140: 683, 1941.
(101) Greenberg, L. J. Pharmacol 77: 194, 1943.
(102) Westerfeld, W. W. and R. L. Berg. J. Biol. Chem. 148: 523, 1943.
(103) Deffner, M. Annalen 536: 44, 1938.
(104) Slade, H. D. and C. H. Werkmann. J. Bact. 41: 675, 1940.
(105) Campbell, J. J. R. and I. C. Gunsalus. J. Bact. 48: 71, 1944.

(106) GUNSALUS, I. C. AND J. J. R. CAMPBELL. J. Bact. 48: 455, 1944.
(107) LIPTON, M. A. AND E. S. G. BARRON. J. Biol. Chem. 166: 367, 1946.
(108) SLADE, H. D. AND C. H. WERKMANN. Arch. Biochem. 2: 97, 1943.
(109) ROBINSON, R. J. Soc. Chem. Ind. 53: 1062, 1934.
(110) FRASER, M. T. AND J. A. GARDNER. Proc. Roy. Soc. B. 82: 55, 1910.
(111) WINDAUS, A. Cited in H. LETTRÉ AND H. H. INHOFFEN. Ueber Sterine, Gallensäuren und Verwandte Naturstoffe. Stuttgart 1936, p. 103.
(112) RITTENBERG, D. AND R. SCHOENHEIMER. J. Biol. Chem. 121: 235, 1937.
(113) SHEMIN, D. AND D. RITTENBERG. J. Biol. Chem. 159: 567, 1945.
(114) SHEMIN, D. AND D. RITTENBERG. J. Biol. Chem. 166: 621, 1946.
(115) WOOD, H. G., N. LIFSON AND V. LORBER. J. Biol. Chem. 159: 475, 1945.
(116) STETTEN, D., JR. AND M. R. STETTEN. J. Biol. Chem. 165: 147, 1946.
(117) RITTENBERG, D. AND K. BLOCH. J. Biol. Chem. 154: 311, 1944; 160: 417, 1945.
(118) BLOCH, K. AND D. RITTENBERG. J. Biol. Chem. 145: 625, 1942.
(119) DEUEL, H. J., JR. AND A. T. MILHORAT. J. Biol. Chem. 78: 299, 1928.
(120) BUCHANAN, J. M., A. B. HASTINGS AND F. B. NESBETT. J. Biol. Chem. 150: 413, 1943.
(121) LORBER, V., N. LIFSON AND H. G. WOOD. J. Biol. Chem. 161: 411, 1945.
(122) BUCHANAN, J. M. AND A. B. HASTINGS. Physiol. Rev. 26: 120, 1946.
(123) WOOD, H. G. Physiol. Rev. 26: 198, 1946.
(124) EMBDEN, G. AND F. OPPENHEIMER. Biochem. Ztschr. 45: 186, 1912.
(125) LEHNINGER, A. L. J. Biol. Chem. 164: 291, 1946.
(126) MACLEAN, I. S. AND D. HOFFERT. Biochem. J. 20: 343, 1926.
(127) SONDERHOFF, R. AND H. THOMAS. Annalen 530: 195, 1937.
(128) WOOD, H. G., R. W. BROWN AND C. H. WERKMANN. Arch. Biochem. 6: 243, 1945.
(129) BARKER, H. A., M. D. KAMEN AND B. T. BORNSTEIN. Proc. Natl. Ac. Sci. 31: 373, 1945.
(130) STETTEN, D., JR. AND R. SCHOENHEIMER. J. Biol. Chem. 133: 329, 1940.
(131) BLOCH, K., E. BOREK AND D. RITTENBERG. J. Biol. Chem. 162: 441, 1946.
(132) STETTEN, D., JR. AND G. E. BOXER. J. Biol. Chem. 156: 271, 1944.
(133) BLOCH, K. AND D. RITTENBERG. J. Biol. Chem. 143: 297, 1942.
(134) BLOCH, K., B. N. BERG AND D. RITTENBERG. J. Biol. Chem. 149: 51, 1943.
(135) BLOCH, K. J. Biol. Chem. 157: 661, 1945.
(136) BLOCH, K. AND D. RITTENBERG. Unpublished.
(137) SHEMIN, D. AND D. RITTENBERG. J. Biol. Chem. 166: 627, 1946.
(138) FISCHER, H. AND E. FINK. Ztschr. physiol. Chem. 280: 123, 1944.
(139) KNOOP, L. Ber. Chem. Ges. 17: 1638, 1884.
(140) BARKER, H. A. AND J. V. BECK. J. Biol. Chem. 141: 3, 1941.
(141) ALMQUIST, H. J., E. L. R. STOKSTAD, E. MECCHI AND P. D. V. MANNING. J. Biol. Chem. 134: 213, 1940.
(142) GUIRARD, B. M., E. E. SNELL AND R. J. WILLIAMS. Arch. Biochem. 9: 361, 1946.
(143) SCHOENHEIMER, R. The dynamic state of body constituents. Harvard University Press, Cambridge, 1942.
(144) KHARASCH, M. S. AND M. T. GLADSTONE. J. Am. Chem. Soc. 65: 15, 1943.
(145) UTTER, M. F., F. LIPMANN AND C. H. WERKMANN. J. Biol. Chem. 158: 521, 1945.
(146) DU VIGNEAUD, V. The Harvey Lectures, Series 38: p. 39, 1942.
(147) BORSOOK, H. AND J. W. DUBNOFF. Science 91: 551 1940.
(148) BLOCH, K. AND R. SCHOENHEIMER. J. Biol. Chem. 134: 785 1940.
(149) SOLOMON, A. K., B. VENNESLAND, F. W. KLEMPERER, J. M. BUCHANAN AND A. B. HASTINGS. J. Biol. Chem. 140: 171, 1941.
(150) ELLIOT, K. A. C., M. B. BENOY AND Z. BAKER. Biochem. J. 29: 1936, 1935.
(151) TOENNIESSEN, E. AND E. BRINKMANN. Ztschr. physiol. Chem. 252: 169, 1938.
(152) KLEINZELLER, A. Biochem. J. 37: 674, 1944.
(153) WIELAND, H. AND R. SONDERHOFF. Annalen 499: 213, 1932.

(154) LIPMANN, F. Advances in Enzymology, New York 1: 99, 1941.
(155) WIELAND, H., G. JENNEN AND W. SCHWARZE. Annalen 548: 255, 1941.
(156) JOWETT, M. AND J. H. QUASTEL. Biochem. J. 29: 2181, 1935.
(157) LYNEN, F. Annalen 554: 40, 1943.
(158) BREUSCH, F. L. Science 97: 490, 1943.
(159) WEIL-MALHERBE, H. Nature 153: 435, 1944.
(160) KREBS, H. A. AND L. EGGLESTON. Nature 154: 209, 1944.
(161) WERKMANN, C. H. AND H. G. WOOD. Advances in Enzymology, New York 2: 135, 1942.
(162) KREBS, H. A. Advances in Enzymology, New York 3: 191, 1943.
(163) EVANS, E. A., JR. The Harvey Lecture, Series 39: 273, 1944.
(164) LIPMANN, F. Advances in Enzymology, New York 6: 231, 1946.
(165) LIPMANN, F. J. Biol. Chem. 155: 55, 1944.
(166) LIPMANN, F. AND C. L. TUTTLE. J. Biol. Chem. 161: 415, 1945.
(167) BUCHANAN, J. M. AND J. C. SONNE. J. Biol. Chem. 166: 781, 1946.
(168) OCHOA, S. J. Biol. Chem. 159: 243, 1945.
(169) KREBS, H. A. AND W. A. JOHNSON. Biochem. J. 31: 772, 1937.
(170) LEHNINGER, A. L. J. Biol. Chem. 148: 393, 1943.
(171) KREBS, H. A. Nature 138: 288, 1936.
(172) WEINHOUSE, S., G. MEDES, N. FLOYD AND L. NODA. J. Biol. Chem. 161: 745, 1945.
(173) MARTIUS, C. Ztschr. Physiol. Chem. 257: 29, 1938.
(174) KREBS, H. A. Biochem. J. 36: ix, 1942.
(175) WEINHOUSE, S., G. MEDES AND N. FLOYD. J. Biol. Chem. 166: 691, 1946.
(176) WIELAND, H. AND C. ROSENTHAL. Annalen 554: 241, 1943.
(177) HUNTER, E. AND L. F. LELOIR. J. Biol. Chem. 159: 295, 1945.
(178) VIRTANEN, A. J. Ann. Ac. Sc. Fenn. 33: 3, 1930.
(179) SONDERHOFF, R. AND M. DEFFNER. Annalen 530: 195, 1937.
(180) BUCHANAN, J. M., W. SAKAMI, S. GURIN AND D. W. WILSON. J. Biol. Chem. 159: 695, 1945.
(181) BREUSCH, F. L. Enzymologia 11: 169, 1944.
(182) WOOD, H. G., N. LIFSON AND V. LORBER. Unpublished results, cited in Physiol. Rev. 26: 198, 1946.
(183) RITTENBERG, D. AND K. BLOCH. J. Biol. Chem. 157: 749, 1945.
(184) BUCHANAN, J., W. SAKAMI, S. GURIN AND D. W. WILSON. Fed Proc. 5: 126, 1946.
(185) KREBS, H. A. AND W. A. JOHNSON. Enzymologia 4: 148, 1937.
(186) EVANS, E. A., JR. AND L. SLOTIN. J. Biol. Chem. 141: 439, 1941.
(187) WOOD, H. G., C. H. WERKMANN, A. HEMINGWAY AND A. O. NIER. J. Biol. Chem. 142: 31, 1942.
(188) LYNEN, F. Annalen 552: 272, 1942.
(189) MARTIUS, C. Ztschr. physiol. Chem. 274: 96, 1943.
(190) HAHN, L., G. HEVESY AND E. LUNDSGAARD. Biochem. J. 31: 1705, 1937.
(191) BLIXENKRONE-MOELLER, N. Ztschr. physiol. Chem. 252: 137, 1938.

BIOLOGICAL AND MEDICAL APPLICATIONS OF ELECTROPHORESIS

JOHN A. LUETSCHER, Jr.

Department of Medicine, Johns Hopkins University and Hospital, Baltimore, Md.

The need of an effective method for the separation and analysis of mixed proteins has been uniquely met by recent improvements in the electrophoretic method. The ordinary principles of qualitative analysis are inadequate to deal with proteins, whose components are similar and whose characteristics are determined less by their composition than by their structure. Furthermore, proteins may be easily modified and their distinctive properties may be destroyed by the chemical treatment necessary for their separation. The physical methods are better fitted for the task. A distinctive structural characteristic of the protein is used in the separation, and the forces applied are extraordinarily harmless. The amount of material necessary for physical analysis is small. Ultracentrifugation (329) and electrophoresis share these properties and yield valuable, independent information on the characteristics and composition of proteins.

Technique. The movement of protein ions in an electric field has emerged during the last decade as an important biochemical method (45, 165, 170, 186, 214, 323, 324, 328, 338, 355, 359). Earlier studies were limited by technical difficulties and the present usefulness of the electrophoretic method has developed largely from improvements in technique. Ten years ago, Tiselius described the apparatus which gave speed and precision to the moving boundary method and made possible the quantitative study of complex proteins (348). Longsworth, Philpot, and Svensson adapted the optical system to simplify the quantitative analysis of mixtures (116, 131, 163, 164, 165, 170, 173, 213, 250, 330, 331, 338). These technical improvements opened new fields of investigation, which have been widely explored and described in a literature so extensive as to be difficult to bring together in a review of reasonable length. This review is limited almost entirely to the work of the last decade. The reader will doubtless find applications to his own field of interest which can be fully appreciated only after study of original work in that field.

In the "moving-boundary" method of electrophoresis, a sharp "boundary" between a protein solution and its solvent is formed and observed as the protein moves in an electric field. If the protein is homogeneous in mobility, the boundary remains sharp and single. If there are different species of protein in the solution, they may move with different speeds, so that the original, sharp, protein-solvent boundary becomes wider and may separate into several boundaries, each representing the end of a column of protein of different mobility. The number and position of the boundaries are observed by means of the change in refractive index produced by the change in protein concentration. If the specific refractive increment of each protein is known, the concentration of the components of a mixture may be measured. Under suitable conditions, the

slowest and fastest components may be separated from the mixture and recovered in nearly pure form.

Electrophoretic study of insoluble proteins can be made by the microscopic observation of the motion of particles. Inert particles coated with protein generally move with the same mobility as the dissolved protein (221, 222, 223, 300, 307). The surface of living cells may be studied by similar observations. Abramson, Moyer, and Gorin (10) have reviewed the earlier micro-electrophoretic studies and included an account of their interesting studies on cells. Electro-osmosis and membrane potentials are considered to be corollaries of electrophoretic migration.

Observation of purified proteins has brought out the dependence of electrophoretic mobility on the pH and ionic strength of the solution, and curves of mobility at various hydrogen-ion concentrations are closely related to the acid-base titration curves of the protein or other amphoteric substance (4, 10, 54, 322, 338, 360). The mobility of protein is affected not only by the quantity of ions present but also by the types of ion in the solvent (16, 338, 349). Since the relative mobilities of the components of a mixture determine the completeness of electrophoretic separation, the choice of buffer is of critical importance. Furthermore, the boundary anomalies and the corresponding uncertainty of analytical accuracy may be reduced in proper buffers (66, 165, 167, 170, 336, 338). The avoidance of convection requires accurate adjustment to the temperature of maximum density of the solution studied (12, 165, 207, 211).

Electrophoretic Observations and Applications. The simplest type of observation which can be made by electrophoresis is the enumeration of boundaries. The component protein fractions may be further described according to their mobility under one or more conditions of pH and ionic strength. Tiselius and Stenhagen described the electrophoretic fractions of the plasma proteins in this way, identifying albumin, fibrinogen, and the three globulin fractions which were named alpha, beta, and gamma in the order of their mobilities (320, 348, 349, 350).

Electrophoretic homogeneity is sometimes interpreted as chemical homogeneity without adequate proof or reservation. There are many other important criteria of the purity of proteins, including chemical analysis, crystallinity, ultracentrifugal analysis, and, perhaps, most important, constant solubility (302). Biological activity may be used as a crude measure when present. These methods are unfortunately contradictory in some cases. The weight which is placed on the electrophoretic measurements should depend on the completeness of examination under wide variations of pH, ionic strength, and buffer salts, with a protein concentration adequate to detect small amounts of an impurity. In certain cases, it is easy to decide the relative value of electrophoresis. Crystals of serum albumin, egg albumin, or hemoglobin, studied by electrophoresis, may show recognizable contaminants from the native mixture (145, 168; 187, 216). A pure preparation may become electrophoretically impure on standing (216, 304). On the other hand, insulin appears the same to electrophoresis, whether the preparation is crude or purified (100), and solubility has proved a

more sensitive criterion than electrophoresis in the study of pepsin (107) and chymotrypsinogen (28). Ultracentrifugation may yield simpler patterns than electrophoresis, as in normal serum globulin (242, 329) and enzymatic digests (61, 248, 362), or more complex patterns than electrophoresis, as in the abnormal serum globulins of multiple myeloma (130, 196, 242). Electrophoretic homogeneity is obviously a very limited measure of purity when the preparation has been made by electrophoretic separation.

The most fruitful use of this type of observation has been in the guidance of chemical separation of proteins from natural sources. The results have been equally important to the student of electrophoresis, since they have identified the chemical and biological nature of the electrophoretic components.

In extensive studies of plasma proteins, E. J. Cohn and his associates have used the electrophoretic method as a control in fractionation (43, 44, 45, 46), obtaining not only purified fractions of biological and medical value, but also a vast store of information correlating chemical and biological properties with electrophoretic behavior. Electrophoretic control of classical and newer methods of plasma fractionation has found general acceptance (61, 62, 63, 93, 119, 332, 349, 350, 381). This co-ordinated approach has been less often applied to the proteins altered by disease in order to elucidate the changes inadequately described by electrophoresis alone. Longsworth (169) demonstrated the increased lipoid content of the beta-globulin in nephrosis and in obstructive jaundice. Recently, Seibert (290) has correlated the increased polysaccharide content of blood with the increase in alpha-2-globulin seen in pneumonia, tuberculosis, and certain forms of cancer.

The striking changes which occur in the electrophoretic pattern of serum during immunization and after reaction with the specific antigen give a clear picture of the appearance and removal of antibody. Such studies have been equally important in the understanding of the electrophoretic pattern and in establishing the relationship of antibody to the plasma proteins. The early studies of Tiselius and Kabat showed that circulating antibodies appeared in the gamma-globulin fraction of rabbit plasma (350, 353, 358) and this has been abundantly confirmed in many species (52, 53, 57, 73, 132, 206, 209, 231, 232, 274, 289, 341, 345, 365, 366, 369, 382, 383). Even the "normal" adult man bears the scars of old battles with disease in his gamma-globulin (63, 71). The increased globulin after immunization is not necessarily identical with normal gamma-globulin (63, 104, 174, 353, 358, 365, 366, 368), nor is the increase in the gamma fraction entirely active antibody (270, 368, 371, 382). Less often, some antibody has been found in other globulin fractions (71, 73, 132, 239, 286). The electrophoretic method has been useful in studying the chemical purification of immune globulin, but has found only a minor place in the preparation of highly purified antibody (335), since other methods have proved to be more specific and effective.

Some information on the origin of antibody has been contributed by White and Dougherty (69, 378, 379), who have shown that several proteins resembling those of serum may be formed in lymph nodes. These proteins are released from

the lymph nodes under pituitary adrenocortical stimulation at the same time that antibody increases in the serum. Since the adrenocorticotrophic hormone may also increase serum albumin (160), the ensuing relative changes in serum proteins may vary and can not be interpreted in the absence of measurements of plasma volume. It seems quite clear, however, that several proteins, including antibody, may be increased in the circulating blood after pituitary adrenocortical stimulation. The rôle of the lymph node in the production of gamma-globulin has been corroborated by the demonstration in lymph nodes of a protein antigenically identical with the gamma-globulin of serum (127).

The purification of proteins with hormonal and enzymatic activity has been aided by electrophoresis as a control and, in some cases, as a separative tool. Theorell combined the chemical and electrophoretic methods in his purification of the yellow enzyme (128, 343) and later of cytochrome C (344). Preparations of pepsin (106, 107, 352), chymotrypsinogen (28), phosphorylase (89, 374), zymohexase (310), 1-amino-oxidase (90), carboxylase (202), ribonuclease (269), mucopolysaccharase (98, 99), and carbonic anhydrase (247, 311) have been examined. Enzymatic digestion of protein has been followed by electrophoresis (61, 248, 249, 356, 362, 363, 370), but ultracentrifugation is generally a more sensitive measure. The preparations of the protein hormones of the pituitary have been guided by electrophoresis (38, 39, 151, 152, 153, 154, 156, 157, 159, 228, 277, 299, 304, 376, 377). In the case of insulin, however, the method has not proved applicable (100). Highly purified thyroglobulin is homogeneous in its migration (103).

Many proteins of cells and their secretions (35, 36, 58, 59, 70, 91, 101, 114, 115, 118, 122, 133, 155, 177, 200, 201, 217, 225, 258, 265, 315, 325, 327, 354, 357, 380, 387, 388) have been subjected to electrophoresis, generally to estimate the purity or complexity of a preparation or to separate an active material, such as the hematopoietic factor of the liver (375) or the thromboplastin of lung (41). Other materials, such as hemoglobin or egg albumin (11, 13, 18, 29, 54, 83, 168, 240, 263) have been extensively studied because of their availability, ready crystallization, and the wealth of comparative physico-chemical information, as well as for the interest in their biological properties (27). Bacterial antigens and toxins (1, 47, 82, 125, 126, 142, 162, 192), as well as active preparations of virus (24, 25, 72, 79, 146, 197, 198, 204, 229, 256, 257, 271, 293, 300, 303, 307, 308) have been studied. Plant extracts of many types show complex patterns which can usually be related to their chemically separated components (15, 75, 76, 77, 117, 264, 273, 280).

A number of interesting investigations deal with cell proteins which commonly give rise to hypersensitivity in man. Seibert's purification of tuberculin proteins, guided by electrophoresis, has led to a standardized testing material of medical importance (92, 282, 283, 284, 285). Abramson, Moore, and Gettner have separated the proteins of hay-fever producing pollens by electrophoresis and have identified skin-reactive fractions of low molecular weight, which may account for the unusual sensitizing ability of these pollens (6, 7, 8, 9). Chemical methods of separation of pollen fractions have not been so effective under electro-

phoretic scrutiny (233). Electrophoresis has found a place in the separation of the allergenic protein of cottonseed (316, 317).

Intact living cells move in the electric field. The influence of the solvent on this motion may reveal properties of cell surfaces, which determine the mobility (2, 10, 80, 105, 319). Intact red blood cells possess distinctive surface characteristics which are little altered by suspension in protein solutions (5) and surprisingly unaffected by damaging agents (14). The ghosts of hemolyzed red blood cells migrate with the same speed as the intact cell (3) and a component of similar mobility has been identified in extracts of the cells (326). Spermatozoa show variations in isoelectric point and mobility, which allow separation and perhaps some degree of sex differentiation (305, 306). Bacteria migrate in the electric field, and some biological characteristics may be correlated with mobility (26, 33, 78, 161, 236, 319, 372). Similar studies of protozoa have been made (193, 291, 305).

Preparative Electrophoresis. It is often desirable to recover one or more of the fractions separated by electrophoresis for chemical characterization or for study of biological activity. Since separation and recovery are most successful when the mobilities of the components are distinctly different, the conditions should be chosen with great care. In any single separation, only the fastest and slowest components can be obtained in pure form. Fractions of intermediate speed can be obtained by two or more steps, first isolating two faster components, for example, and then separating them in a second operation. Given a mixture which can theoretically be separated, the chief problems are concerned with technique, purity, and yield. In general, techniques which give fractions of the highest purity also yield very little material (337) and the collection of large amounts of material is very tedious. The conventional U-tube design of Theorell (342) and Tiselius (348, 351) includes apparatus of varying size and with varying numbers of compartments which can be closed off at the end of the separation. Macheboeuf (184) described an ingenious method of withdrawing samples without disassembling the apparatus. Svensson (333, 334, 338) uses a similar technique in his apparatus, which is arranged to allow almost continuous operation. There are other quite different principles of separation which have not yet reached the technical perfection of the conventional apparatus (337, 338), but which have great promise in suitable circumstances. "Electrophoresis-convection" causes the concentration of protein against a semi-permeable membrane (94, 135, 343). Stationary electrolysis separates proteins of different isoelectric points (97, 361). Flowing, thin layers of protein have been used to facilitate separation (251). Packchanian described a large horizontal apparatus suitable for separation of living organisms (238).

Analysis of Plasma Proteins. The quantitative analysis of the component proteins of mixtures has had its widest application in the analysis of plasma and of certain related fluids. The original nomenclature of Tiselius is still in general use for the serum globulins, which are described as alpha, beta and gamma in the order of their mobilities in alkaline solutions. Certain subdivisions have appeared: two alpha and two beta globulin fractions are often further identified by

the subscripts 1 and 2. The complexity of the fibrinogen fraction is indicated by the appreciable protein of the same mobility remaining after clotting (385), which has been isolated and found to resemble gamma-globulin (63). The wide, reversible spreading of the gamma boundary shows its heterogeneity (116, 121, 294, 295). Albumin separates into two components on prolonged electrophoresis (22, 113), especially just above pH 4 (178, 296) or in the presence of certain anions (16, 336).

In the "normal" man and in certain animals, the electrophoretic fractions have acquired more or less substance of chemical and physical characterization. Such information confirms the complex nature of the albumin and globulin fractions (43, 44, 45, 46, 112, 199, 296, 332), as well as the presence of biologically active proteins in amounts too small to recognize in the pattern of the native proteins (45, 237, 252, 255, 281, 321). The association of carbohydrates and lipoids with the blood proteins is very intimate (21, 22, 34, 185, 350, 386) and these substances affect both the refractive increment and mobility of the proteins to which they are bound, chiefly the alpha- and beta-globulins (21, 23, 140, 143, 249, 385, 386).

When the proportions of the serum proteins are altered by disease, there may be qualitative changes in the fractions and not simply more or less of the normal proteins. In the study of such changes, it is important to employ suitable physical, chemical and biological techniques to correlate the electrophoretic patterns with more definitive alterations in the proteins. Many current reports are descriptive, a natural stage in the collection of information. It is to be hoped that the interest so aroused will stimulate the collection of more fundamental information on the altered proteins, so that these fractions, now only shadows on the electrophoretic diagrams, may be better defined.

The electrophoretic analysis of human plasma has become rigidly standardized in many laboratories, since not only the quantitative distribution of proteins but even the identity of fractions may vary somewhat with changes in the buffer (16, 50, 136, 170, 217, 245, 336). The human alpha-1-"globulin" of the most generally used barbiturate buffer is included in the "albumin" fraction in phosphate buffer, while the alpha-2 and beta globulins are often more clearly separated in phosphate buffer. Two useful buffers about which the most information has accumulated are sodium diethyl barbiturate of ionic strength 0.1 and pH 8.6, and sodium or potassium phosphate of ionic strength 0.2 and pH 7.7. Since too high a protein concentration affects the analytical results in buffer of low ionic strength (16, 65, 345, 336), the design of the optical system must conform to certain standards in order to obtain the best results. This total standardization leads to a single end, the comparability of human plasma patterns. For purposes other than the standard quantitative analysis of human serum or plasma, the investigator can adapt these conditions to produce optimum separation of the components which he wishes to study. The enormous flexibility of the electrophoretic technique should not be impaired by too rigid conventions, nor should sight be lost of the artificiality of any such single set of conditions.

The interpretation of the shadow diagrams must be guarded in certain respects. Interaction between proteins and other proteins (35, 47, 139, 141, 168, 217, 354,

357, 359) or smaller ions (16, 17, 171, 183, 205, 230, 259, 261, 322) may produce anomalous results. The changes occurring after prolonged standing (16, 143, 176, 187, 216, 279), freezing, drying, or denaturation (49, 55, 64, 102, 138, 140, 141, 188, 248, 262, 309, 350, 367) are defined for some proteins and mixtures. Moreover, the usual analytical method assumes the uniform refractive increment of components. These include lipoids and carbohydrates bound to serum proteins, and conversion factors are necessary in order to apply the data to nitrogen content (16, 48). The normal conversion factors may change under pathological conditions (16, 169, 386).

In spite of these criticisms, the electrophoretic method gives an accurate, reproducible, and distinctive description of the plasma proteins, which is not attained by any other available method (16, 50, 338, 354). It was long ago recognized that the "salting-out" of proteins yields only a limited analytical approach (37, 65, 88, 179, 219, 341, 349, 350, 371, 385). Original globulin precipitates studied by electrophoresis are complex mixtures, and the separation of albumin from globulin requires several recrystallizations. Alcohols yield better separations (44, 46, 254, 332), but require exact control of conditions.

The normal pattern of plasma has been established not only for man (16, 65, 88, 95, 129, 143, 144, 166, 179, 210, 235, 245, 254, 272, 279, 301, 330, 341), but for a number of animals (60, 136, 150, 158, 218, 275, 297, 336, 338). The analyses vary somewhat with different buffers (218, 336, 338) but indicate a striking species variation in the number and distribution of protein components. The normal electrophoretic pattern of human plasma is comparatively simple, since the separation of albumin, fibrinogen, and the three globulins is quite complete. The nomenclature of Tiselius for these well-defined fractions becomes quite artificial when applied to very complex animal sera, such as those of the sheep, cat, and guinea pig, in which as many as ten peaks may appear (60). Arbitrary definition is certain to lead to confusion, and a surer method of description of such complex sera is by mobility under defined conditions of buffer ions, ionic strength, pH, and protein concentration. The buffers in which many of these sera have been examined were chosen by systematic trial for optimal separation of human serum. More satisfactory buffers may well be found for individual animal sera.

From reported analyses it appears that ungulates tend to have a lower proportion of albumin in plasma than rodents and primates (60, 218, 338). The horse has the lowest recorded proportion of albumin. The plasma proteins of the rhesus monkey have a pattern closely resembling the human pattern. Birds show a rather characteristic mass of slower-moving globulins; and the "beta-globulin disturbance," fibrinogen, and "gamma-globulin" are poorly separated. The hen and cock may show differences in serum proteins, while in mammals the serum proteins show little or no sex differentiation. The carp has a pattern quite different from the higher animals (60, 218). Animals show considerable variations of protein distribution between individuals and breeds in addition to the species differences (218, 297, 298, 338).

In developing chick and pig embryos, the plasma undergoes a gradual differen-

tiation from a simple to a more complex pattern (220). The carbohydrate content falls during the period of development, while the nitrogen content rises. Albumin appears later than several components with the mobility of globulin (220). The blood proteins of the ruminant fetus are apparently quite different from those of the human or rodent fetus. The human fetus at birth has only a slightly lower albumin and higher gamma globulin than the adult man (144, 172, 279). The new-born calf, however, has almost no gamma-globulin and little beta-globulin until colostrum is ingested, when there is a remarkable increase in these globulins and the associated antibodies (120, 274). Pedersen (241, 242) has found a characteristic serum globulin which comprises a large part of the globulin of the ruminant fetus, but is found only in small amounts in the fetus of man or rabbit. These differences are probably related to the less permeable placenta of the ruminant. After weaning, the young dog may develop a lower albumin, fibrinogen, and gamma-globulin, and a higher alpha- and beta-globulin than the adult dog (150).

When quantitative analysis by electrophoresis was applied to human plasma in disease, surveys indicated such distinctive patterns as to arouse hopes for a wide application in medical diagnosis. Characteristic electrophoretic patterns have indeed become recognized as virtually diagnostic of certain diseases, such as nephrosis and some types of myelomatosis. More often, however, the picture has proved to be characteristic not of the specific disease but of the host's reaction to infection or injury. The various changes are frequently proportional to the severity of the physiological disturbance and may vary with the duration or stage of the disease, with nutritional factors, with loss of plasma protein, and with the involvement of certain organs, such as the liver (88, 182, 218, 290).

The common denominator of almost every pathological state is a relative or absolute decrease in the serum albumin. The electrophoretic technique demonstrates this change with greater sensitivity than the salting-out methods. The reduction of serum albumin is often associated with two factors common to many diseases,—a deficiency of protein and a general reaction of the body to injury and infection (31, 81, 190, 244). When the concentration of serum albumin is seriously reduced, the concentration is stabilized by osmotic factors and by the removal of transfused albumin from the circulation (181). Electrophoretically homogeneous fractions, such as concentrated human serum albumin, can be effectively used to study such transfers of protein (181, 346, 347).

The effects of malnutrition on the electrophoretic pattern of human plasma have not been adequately described. To be sure, there seem to be as many possible types of malnutrition as types of disease, but very little fundamental information of any sort is available. Keys (134) found normal electrophoretic patterns in volunteers who showed great weight loss and edema on a diet characteristic of the recent European famine. Patients with a history of prolonged alcoholism and multiple dietary deficiencies have developed intractable hypoproteinemia with a reduction of both serum albumin and globulin (182), but similar changes have been observed in patients without obvious dietary deficien-

cies (328). In one case, a very low concentration of gamma-globulin has been observed to increase after an adequate intake of protein (137).

The plasma of dogs on a diet deficient in protein with or without plasmapheresis shows a considerable reduction in serum albumin (37, 384, 385). The globulins are apparently increased, due to associated lipids, but the globulin-nitrogen of plasma is not increased (385, 386). The regeneration of albumin after depletion by diet and plasmapheresis takes a matter of weeks and is still incomplete when the total plasma protein has returned to normal levels (384, 385). Pigs on a tryptophane-deficient diet may develop a marked reduction in serum albumin and in several globulin fractions (30).

Severe bodily injury is followed by striking changes in the albumin and alpha-globulin fractions of the plasma. Burns, freezing, skin irritants, certain poisons, hemorrhage, fractures of bones, and myocardial infarction are followed by an increased alpha-globulin concentration (31, 81, 244, 301). The reduction of albumin may be a measure of the severity of injury and occurs earlier than the increase in globulin. The close resemblance to the changes observed in acute, febrile infections is unmistakable and suggests that this reaction of the proteins to injury is related to adrenal stimulation and the protein-catabolic reaction, which appear in similar circumstances.

The plasma globulins respond to infection in two distinctive patterns. In the febrile stage of an acute infection there is usually an increase of alpha-globulin to several times its normal concentration. The usual reduction in serum albumin makes the proportional rise of the alpha-globulin even more striking. This change has been observed in pneumonia (166, 180), tonsillitis (166, 190), scarlet fever (68), rheumatic fever (68, 166, 190), peritonitis (166), and infectious mononucleosis (42). An exception to the general association of acute febrile illness with an increased alpha-globulin occurs in vivax malaria (67, 96). The increased alpha-globulin may be related to the "C-reactive protein" which appears in human serum in acute infections and which is found in the alpha-globulin fraction (243).

A later change in serum globulin following virtually all types of infection is an increased concentration of the gamma-fraction (42, 53, 67, 68, 96, 166, 180, 190, 191, 208, 231, 272, 288, 290). The time-relationship of these changes has been clearly demonstrated in a continuous study of the plasma during the course of scarlet fever and rheumatic fever (68). The increase in gamma-globulin is probably correlated with the production of antibodies. The new component may not have a mobility identical with the normal fraction (63). In kala-azar, a fraction of unusually slow mobility appears in large amounts (51).

In chronic infections of many types, there is a tendency for all globulins to become increased in a manner reminiscent of "hyperimmune" sera of animals (366). In many infections, fibrinogen is increased at an early stage along with alpha globulin and may remain elevated if the disease persists (190). A recent survey of many cases of tuberculosis describes changes in the serum globulins varying with the stage of the disease (290). In minimal tuberculosis, gamma globulin is elevated above the normal level. Moderately advanced cases show

parallel increases in the alpha-2-globulin fraction and in the closely associated polysaccharide of the serum. In far advanced tuberculosis, all globulin fractions are increased (287, 288, 290). The total protein of the serum usually remains normal in spite of a progressive fall in the albumin concentration (290). The striking changes in the alpha-2-globulin in tuberculosis have been variously thought to be associated with a specific antibody to a fraction of tuberculin (286, 289), with a sensitization of the host (286), and with tissue destruction (290). A protective antibody against tuberculosis has been identified in gamma-globulin. In sarcoidosis, the hyperproteinemia is frequently due to a large increase in gamma-globulin (74, 288, 290). The other globulins are usually more concentrated than normal, and the polysaccharide associated with the alpha-2-globulin is increased, though not so much as in tuberculosis (290). Leprosy may produce large increases in alpha- and gamma-globulin (288). In syphilis, a small increase in serum globulin fractions, especially gamma (53, 231), is observed. The antibody responsible for the Wassermann and Kahn tests for syphilis is concentrated in a fraction of gamma-globulin (52, 53, 57, 231), but material giving a positive Hinton test has been found in another fraction (52). Anticomplementary activity is a normal property of gamma-globulin, but is inhibited by other serum fractions (56, 235). The hyperglobulinemia of lymphopathia venereum involves all fractions (166).

The common allergic diseases produce little change in the distribution of blood proteins, although antibodies can be identified in the gamma-globulin fraction (232). In the group of diffuse diseases of mesenchymal tissue which partake both of infection and allergy, the globulins show well defined changes. In acute rheumatic fever, there is an increase in alpha-globulin, and probably a larger and more prolonged rise in gamma-globulin than would have occurred in an uncomplicated streptococcal infection (68, 166, 190, 272). Essentially the same changes are seen in rheumatoid arthritis, regressing toward normal in remission or recovery (175, 190, 246, 339, 373). The increased globulin observed regularly in disseminated lupus erythematosus is found in the gamma-fraction (40, 190). Erythema nodosum is accompanied by increased alpha- and gamma-globulin and fibrinogen (190).

The different stages of glomerular nephritis lead to characteristic changes in the plasma proteins. In acute nephritis, albumin is reduced even more than indicated by the conventional salting-out methods (182). Gamma-globulin is increased in both relative and actual concentration and may remain so for many weeks. Alpha- and beta-globulin concentrations are normal unless an active infection is present (182). The nephrotic stage is associated with a profound reduction in the serum albumin and in gamma-globulin, while the alpha- and beta-globulins are increased (166, 169, 179, 190). The pattern is virtually diagnostic of the disease. A considerable part of the increase in alpha- and beta-globulins is due to lipoids associated with these fractions, which contribute more to the refractive increment than to the nitrogen content (16, 169). These fractions also have an appreciable solubility at the high salt concentrations used to precipitate "globulin." These properties result in unusually large discrepancies

between conventional and electrophoretic analyses (179). The albumin of blood and urine in nephrosis differs from normal in its electrophoretic behavior at pH 4, presumably due to some change in formation, since there is no significant difference between the albumins of nephrotic serum and urine (178). In the terminal stage of glomerular nephritis, the plasma shows a moderate reduction of albumin, and usually some increase in the globulins and fibrinogen (179, 190).

The urinary proteins in nephritis are identical in mobility with the corresponding serum fractions (179). The relative proportions of the various fractions in the urine are determined by two factors, the composition of the plasma and the permeability of the injured nephrons. In the presence of a constant renal lesion, it is possible to modify the excretion of a fraction by changing its concentration in the plasma. A renal "clearance" of that fraction can be calculated which appears to be independent of plasma concentration and urine flow over a limited range (181, 182). Spontaneous changes in the plasma proteins are often reflected in the urinary proteins, which may also be affected by a changing renal lesion. In the different stages of nephritis, the permeability of the kidney varies, especially to albumin, which is most readily excreted in the nephrotic stage (169, 179, 181, 182, 190). Fibrinogen and the lipoprotein complexes are only slightly excreted in any stage of the disease (169, 179, 181, 190). Albumin and alpha-1-globulin thus dominate the urinary pattern in the early nephrotic stage, while the onset of the terminal phase may be heralded by an increased proportion of globulins in the urine (20). In amyloid nephrosis, the plasma may contain large amounts of gamma-globulin which appears in considerable concentration in the urine, together with the other proteins of the serum (179). Other types of proteinuria show less distinctive patterns.

Disease of the liver is associated with characteristic changes in the plasma protein. In advanced portal cirrhosis there is a profound reduction of the serum albumin and a large increase in gamma-globulin, while the alpha- and beta-globulins show smaller increases (88, 179, 180, 190). Injury to the liver due to arsenic may produce similar changes, but beta-globulin may be more prominent (88). Infiltration by metastatic carcinoma produces less striking changes, with more reduction of albumin and less increase in gamma-globulin (88). Infectious hepatitis (catarrhal jaundice) is associated with a reduction in albumin and an increase in gamma-globulin, often with smaller increases in alpha-2- and beta-globulin (88, 191, 208). Obstruction of the bile ducts resulting in jaundice produces little change in the protein pattern for a time, but later the albumin may fall and beta-globulin may increase, possibly as a result of secondary changes in the liver (88, 166, 384).

Recently, there has been some interest in the signs of apparent liver disease which occur in malaria and infectious mononucleosis (42, 96). The similarity of the protein patterns in many chronic infections and in liver disease suggests that liver dysfunction or injury may play a part in the production of these changes in infections, in addition to the more obvious increment of antibody to the gamma-fraction. The colloidal gold and cephalin flocculation tests for liver disease depend on an increase in the gamma-globulin fraction (85, 124, 208), as well as a

decreased concentration and stabilizing power of the albumin fraction (208). Similar changes are found in malaria (96) and may well be present in other infections. The sedimentation rate of human erythrocytes, often used as a measure of the activity of an infection or of necrosis of tissue, is not specific in its response to any single fraction of the globulins or fibrinogen (97, 190, 290, 301) and may also rise with the increased globulins of liver disease and of other non-infectious origins (190).

Several diseases of the blood cells and of their precursors are followed by changes in the plasma proteins. These changes are particularly difficult to evaluate, since certain secondary effects of the disease may play a rôle in the alterations of proteins. The functions of the hematopoietic systems in the formation of plasma globulins (69, 127, 189, 378, 379) may be disturbed, and the liver may be injured by anemia or infiltration. Fever and malnutrition are frequently encountered. Perhaps for these reasons, the changes observed in leukemia (166), Hodgkin's disease (328), and aplastic anemia (166) are not distinctive. The proportions of the plasma proteins are not disturbed in hemophilia (147), but an anticoagulant and a lack of the normal coagulant have been found in the plasma proteins of hemophiliacs (147, 224). The isoagglutinins of normal blood and certain pathological hemagglutinins have been identified in characteristic protein fractions (253, 318).

In multiple myeloma, the electrophoretic pattern of the plasma protein may be distinctive and sometimes virtually diagnostic of this tumor. Ultracentrifugation may show changes even more unusual than those demonstrated by electrophoresis (130, 196, 242). One of several types of protein with the mobility of a beta- or gamma-globulin may be present in the plasma, sometimes in large amounts (19, 95, 130, 166, 190, 212). The changes in the plasma are quite different in individual cases, ranging from a normal pattern or a slight reduction in albumin to a striking hyperproteinemia due to the accumulation of the abnormal protein. The urine may contain Bence-Jones protein of a mobility similar to the abnormal component of plasma, or may be free of protein despite the presence of large amounts of the pathological component in the plasma, or may contain considerable amounts of Bence-Jones protein which does not appear in detectable amounts in the plasma. When the abnormal component has the mobility of a beta globulin, its characteristics in the blood and urine are similar; but when the component of the mobility of gamma-globulin appears in the plasma, the urinary protein is apparently different from the new component of the blood (212). There is some loss of the normal serum proteins in the urine in many cases.

A reduction in serum albumin accompanies many types of tumor. The alpha-2-globulin and an associated polysaccharide are increased in many cases of carcinoma (290). When the liver is involved, the gamma-fraction may become more prominent (88, 290). In a similar way, the changes in other tumors may be more characteristic of the physiological changes than of the tumor itself (328). Infectious tumors of animals may cause changes in the plasma proteins. In certain strains of rabbits infected with papilloma, the beta-globulin fraction and

an associated lipoid are increased (298, 340). Fowl with leucosis show an increase in an abnormal component of the gamma-fraction as early as three days after infection (275). Injection of heat-inactivated tumor produces a similar change which occurs later and is less striking.

Several diseases of the endocrine glands cause changes in the plasma proteins. A reduction in serum albumin occurs in diabetes (149), Addison's disease (194, 195), and both over and under function of the thyroid (148) Corrective treatment is followed by a return of the concentration of albumin toward normal. Desoxycorticosterone acetate may not be as effective as whole adrenal extract in the correction of the depression of albumin in Addison's disease (195). The albumin may remain low despite conventional therapy when certain complications of diabetes (149, 278) and hyperthyroidism (148) are present, leading to the suggestion of high-protein diets under these circumstances. The increase in beta-globulin in diabetes (149, 290) and hypothyroidism (148) may be related to elevated blood lipids. In man, the alpha-globulin is increased in hyperthyroidism and decreased in hypothyroidism (148). A decreased albumin and increased alpha- and beta-globulins are observed in pregnancy (144, 172). The important rôle of the pituitary and adrenal glands in the formation of plasma proteins has been discussed in connection with antibodies. The defect in the serum proteins caused by ablation of the hypophysis (158, 215) can be best remedied by the injection of pituitary adrenocorticotrophic hormone (160). Growth hormone may be quantitatively less effective (160). Removal of the thyroid gland in the rat (219) is followed by a reduction of serum albumin, but the change in alpha-globulin appears superficially to be opposite to the change in human hypothyroidism. It may be pertinent that the affinity of lipoids for the various fractions of serum varies from species to species (60, 218). The increased alpha-globulin in the rat might be the equivalent of the increased beta-globulin in human hypothyroidism.

The proteins of several body fluids have been studied in normal and pathological conditions. The lymph (244), serous effusions (32, 109, 180, 276, 288), joint fluids (111, 246), and cerebro-spinal fluid (123) resemble serum in their protein pattern, unless modified by a local inflammatory process. The colloidal gold reaction, used clinically to estimate the protein distribution in spinal fluid, is produced by gamma-globulin and inhibited by albumin (86, 124), as in the case of similar clinical tests on the blood. The proteins of milk (312, 313), seminal plasma (84, 267, 268), pancreatic juice (226), and the fluids of the eye (110) have their own distinctive patterns.

CONCLUSIONS

Electrophoresis offers a versatile and accurate method for the study, separation, and analysis of proteins. Its applications to some biological and medical problems have been described. The method has aided in the separation of biologically important substances from native mixtures.

Electrophoretic analysis gives a distinctive picture of the composition of the proteins of the body. Such analyses may reflect physiological and chemical

changes in the body. The proteins of the plasma show characteristic patterns in a number of abnormal states. Like many other chemical and physiological measurements in disease, electrophoretic analysis of the plasma can not be expected to make a specific, etiological diagnosis, except in the gradually vanishing instance when only one cause of the physiological disturbance is known. Much more information is needed for the fundamental interpretation of the abnormal changes in the plasma proteins.

REFERENCES

(1) Abrams, A., G. Kegeles and G. A. Hottle. J. Biol. Chem. **164:** 63, 1946.
(2) Abramson, H. A. Ann. N. Y. Acad. Sci. **34:** 121, 1939.
(3) Abramson, H. A., R. F. Furchgott and E. Ponder. J. Gen. Physiol. **22:** 545, 1939.
(4) Abramson, H. A. Trans. Faraday Soc. **36:** 5, 1940.
(5) Abramson, H. A., M. H. Gorin and E. Ponder. Cold Spring Harbor Symposia Quant. Biol. **8:** 72, 1940.
(6) Abramson, H. A., D. H. Moore, H. Gettner, J. Gagarin and L. Jennings. Proc. Soc. Exper. Biol. and Med. **44:** 311, 1940.
(7) Abramson, H. A., M. G. Engel and D. H. Moore. J. Allergy **14:** 65, 1942.
(8) Abramson, H. A., D. H. Moore and H. H. Gettner. J. Phys. Chem. **46:** 192, 1942.
(9) Abramson, H. A., D. H. Moore and H. H. Gettner. J. Phys. Chem. **46:** 1129, 1942.
(10) Abramson, H. A., L. S. Moyer and M. H. Gorin. Electrophoresis of proteins and the chemistry of cell surfaces. Reinhold Publishing Corp., New York, 1942.
(11) Adair, G. S. and M. E. Adair. Trans. Faraday Soc. **36:** 23, 1940.
(12) Alvarez-Tostado, C. J. Biol. Chem. **135:** 799, 1940.
(13) Andersch, M. A., D. A. Wilson and M. L. Menten. J. Biol. Chem. **153:** 301, 1944.
(14) Anderson, C. G. and T. J. Mackie. Brit. J. Exper. Path. **20:** 270, 1939.
(15) Andrews, A. C. J. Am. Chem. Soc. **62:** 942, 1940.
(16) Armstrong, S. H., Jr., M. J. E. Budka and K. C. Morrison. J. Am. Chem. Soc. **69:** 416, 1947.
(17) Ballou, G. A., P. D. Boyer and J. M. Luck. J. Biol. Chem. **159:** 111, 1945.
(18) Beek, J., Jr. and A. M. Sookne. J. Research, Natl. Bur. Standards **23:** 271, 1939.
(19) Blackman, S. S., Jr., W. H. Barker, M. V. Buell and B. D. Davis. J. Clin. Investigation **23:** 163, 1944.
(20) Blackman, S. S., Jr. and B. D. Davis. J. Clin. Investigation **22:** 545, 1943.
(21) Blix, G. Acta Physiol. Scand. **1:** 29, 1940.
(22) Blix, G., A. Tiselius and H. Svensson. J. Biol. Chem. **137:** 485, 1941.
(23) Blix, G. J. Biol. Chem. **137:** 495, 1941.
(24) Bourdillon, J. Proc. Soc. Exper. Biol. and Med. **45:** 679, 1940.
(25) Bourdillon, J. and E. H. Lennette. J. Exper. Med. **72:** 11, 1940.
(26) Bradbury, F. R. and D. O. Jordan. Biochem. J. **36:** 287, 1942.
(27) Broda, E. E. and E. Victor. Biochem. J. **34:** 1501, 1940.
(28) Butler, J. A. V. J. Gen. Physiol. **24:** 189, 1940.
(29) Carpenter, D. C. and F. E. Lovelace. J. Am. Chem. Soc. **65:** 2364, 1943.
(30) Cartwright, G. W., M. M. Wintrobe, W. H. Buschke, R. H. Follis, Jr., A. Suksta and S. Humphreys. J. Clin. Investigation **24:** 268, 1945.
(31) Chanutin, A. and E. C. Gjessing. J. Biol. Chem. **165:** 421, 1946.
(32) Chanutin, A., E. C. Gjessing and S. Ludewig. Proc. Soc. Exper. Biol. Med., in press.
(33) Chapman, G. H. and C. W. Lieb. J. Bact. **37:** 111, 1939.
(34) Chargaff, E. Advances in protein chemistry (M. L. Anson and J. T. Edsall). New York, Academic Press, Inc. **1:** 16, 1944.
(35) Chargaff, E., M. Ziff and D. H. Moore. J. Biol. Chem. **139:** 383, 1941.

(36) Chargaff, E., D. H. Moore and A. Bendich. J. Biol. Chem. **145:** 593, 1942.
(37) Chow, B. F., J. B. Allison, W. H. Cole and R. D. Seeley. Proc. Soc. Exper.Biol. and Med. **60:** 14, 1945.
(38) Chow, B. F., H. B. van Dyke, R. O. Greep, A. Rothen and T. Shedlovsky. Endocrinology **30:** 650, 1942.
(39) Cierezko, L. S. and A. White. Fed. Proc. **1**, pt. 2, 105.
(40) Coburn, A. F. and D. H. Moore. Bull. Johns Hopkins Hosp. **73:** 196, 1943.
(41) Cohen, S. S. and E. Chargaff. J. Biol. Chem. **140:** 689, 1941.
(42) Cohn, C. and B. I. Lidman. J. Clin. Investigation **25:** 145, 1946.
(43) Cohn, E. J., T. L. McMeekin, J. L. Oncley, J. M. Newell and W. L. Hughes. J. Am. Chem. Soc. **62:** 3386, 1940.
(44) Cohn, E. J., J. A. Luetscher, Jr., J. L. Oncley, S. H. Armstrong, Jr. and B. D. Davis. J. Am. Chem. Soc. **62:** 3396, 1940.
(45) Cohn, E. J., J. L. Oncley, L. E. Strong, W. L. Hughes, Jr. and S. H. Armstrong, Jr. J. Clin. Investigation **23:** 417, 1944.
(46) Cohn, E. J., L. E. Strong, W. L. Hughes, Jr., D. J. Mulford, J. N. Ashworth, M. Melin and H. L. Taylor. J. Am. Chem. Soc. **68:** 459, 1946.
(47) Combiesco, C. and E. Sorn. Compt. rend. soc. biol. **132:** 172, 1939.
(48) Cook, R. P. Biochem. J. **40:** 41, 1946.
(49) Cooper, G. R. and H. Neurath. J. Phys. Chem. **47:** 383, 1943.
(50) Cooper, G. R. J. Biol. Chem. **158:** 727, 1945.
(51) Cooper, G. R., C. R. Rein and J. W. Beard. Proc. Soc. Exper. Biol. and Med. **61:** 179, 1946.
(52) Cooper, J. A. Proc. Soc. Exper. Biol. and Med. **57:** 248, 1944.
(53) Cooper, J. A. J. Invest. Dermat. **6:** 109, 1945.
(54) Davis, B. D. and E. J. Cohn. J. Am. Chem. Soc. **61:** 2092, 1939.
(55) Davis, B. D., A. Hollaender and J. P. Greenstein. J. Biol. Chem. **146:** 663, 1942.
(56) Davis, B. D., E. A. Kabat, A. Harris and D. H. Moore. J. Immunol. **49:** 223, 1944.
(57) Davis, B. D., D. H. Moore, E. A. Kabat and A. Harris. J. Immunol. **50:** 1, 1945.
(58) Delsal, J. L., M. Macheboeuf, D. Santenoise and E. Stankoff. Compt. rend. soc. biol. **138:** 51, 1944.
(59) Delsal, J. L., M. Macheboeuf, D. Santenoise and E. Stankoff. Compt. rend. soc. biol. **138:** 435, 1944.
(60) Deutsch, H. F. and M. B. Goodloe. J. Biol. Chem. **161:** 1, 1945.
(61) Deutsch, H. F., M. L. Petermann and J. W. Williams. J. Biol. Chem. **164:** 93, 1946.
(62) Deutsch, H. F., L. J. Gosting, R. A. Alberty and J. W. Williams. J. Biol. Chem. **164:** 109, 1946.
(63) Deutsch, H. F., R. A. Alberty and L. J. Gosting. J. Biol. Chem. **165:** 21, 1946.
(64) Dietz, T. J. J. Franklin Inst. **228:** 396, 1939.
(65) Dole, V. P. and E. Braun. J. Clin. Investigation **23:** 708, 1944.
(66) Dole, V. P. J. Am. Chem. Soc. **67:** 1119, 1945.
(67) Dole, V. P., K. Emerson, Jr. and E. Braun. J. Clin. Investigation **24:** 644, 1945.
(68) Dole, V. P., R. F. Watson, S. Rothbard, E. Braun and K. Winfield. J.Clin. Investigation **24:** 648, 1945.
(68a) Dole, V. P. and S. Rothbard. J. Clin. Investigation **26:** 87, 1947.
(69) Dougherty, T. F., A. White and J. H. Chase. Proc. Soc. Exper. Biol. and Med. **56:** 28, 1944.
(70) Dubuisson, M. and J. Jacob. Rev. can. biol. **4:** 426, 1945.
(71) Enders, J. F. J. Clin. Investigation **23:** 510, 1944.
(72) Eriksson-Quensel, I. B. and T. Svedberg. J. Am. Chem. Soc. **58:** 1863, 1936.
(73) Fell, N., K. G. Stern and R. D. Coghill. J. Immunol. **39:** 223, 1940.
(74) Fisher, A. M. and B. D. Davis. Bull. Johns Hopkins Hosp. **71:** 364, 1942.
(75) Fishman, M. M. and L. S. Moyer. J. Gen. Physiol. **25:** 755, 1942.
(76) Fishman, M. M. and L. S. Moyer. Science **95:** 128, 1942.

(77) Fontaine, T. D., G. W. Irving, Jr. and R. C. Warner. Arch. Biochem. **8:** 239, 1945.
(78) Frampton, V. L. and E. M. Hildebrand. J. Bact. **48:** 537, 1944.
(79) Frampton, V. L. and W. N. Takashi. Arch. Biochem. **4:** 249, 1944.
(80) Furchgott, R. F. and E. Ponder. J. Gen. Physiol. **24:** 447, 1941.
(81) Gjessing, E. C. and A. Chanutin. J. Biol. Chem. **165:** 413, 1946.
(82) Goebel, W. F., T. Shedlovsky, G. I. Lavin and M. H. Adams. J. Biol. Chem. **148:** 1, 1943.
(83) Gralen, N. Biochem. J. **33:** 1907, 1939.
(84) Gray, S. and C. Huggins. Proc. Soc. Exper. Biol. and Med. **50:** 351, 1942.
(85) Gray, S. J. Proc. Soc. Exper. Biol. and Med. **51:** 400, 1942.
(86) Gray, S. J. Proc. Soc. Exper. Biol. and Med. **51:** 401, 1942.
(87) Gray, S. J. and E. B. Mitchell. Proc. Soc. Exper. Biol. and Med. **51:** 403, 1942.
(88) Gray, S. J. and E. S. Guzman-Barron. J. Clin. Investigation **22:** 191, 1943.
(89) Green, A. A. J. Biol. Chem. **158:** 315, 1945.
(90) Green, D. E., D. H. Moore, V. Nocito and S. Ratner. J. Biol. Chem. **156:** 383, 1944.
(91) Gregoire, J., M. Macheboeuf, P. Grabar and P. Lepine. Bull. soc. chim. biol. **26:** 432, 1944.
(92) Grönwall, A. The Svedberg Birthday Volume, 1944.
(93) Grönwall, A. Physiol. Abstracts **21:** 825.
(94) Gutfreund, H. Biochem. J. **37:** 186, 1943.
(95) Gutman, A. B., D. H. Moore, E. B. Gutman, V. McClellan and E. A. Kabat. J. Clin. Investigation **20:** 765, 1941.
(96) Guttman, S. A., H. R. Potter, F. M. Hanger, D. B. Moore, P. S. Pierson and D. H. Moore. J. Clin. Investigation **24:** 296, 1945.
(97) Hahn, L. and A. Tiselius. Biochem. Ztschr. **314:** 336, 1943.
(98) Hahn, L. Arkiv. f. Kemi, mineral., geol., 19A, no. 33, 1945.
(99) Hahn, L. Arkiv. f. Kemi, mineral., geol., 22A, no. 2, 1946.
(100) Hall, J. L. J. Biol. Chem. **139:** 175, 1941.
(101) Hall, J. L. J. Am. Chem. Soc. **63:** 794, 1941.
(102) Hardt, C. R., I. F. Huddleson and C. D. Ball. J. Biol. Chem. **163:** 211, 1946.
(103) Heidelberger, M. and K. O. Pedersen. J. Gen. Physiol. **19:** 95, 1935.
(104) Heidelberger, M., K. O. Pedersen and A. Tiselius. Nature **138:** 165, 1936.
(105) Hennig, K. and H. Ay. Biochem. Ztschr. **299:** 123, 1938.
(106) Herriott, R. M., V. Desreux and J. H. Northrop. J. Gen. Physiol. **23:** 439, 1940.
(107) Herriott, R. M., V. Desreux and J. H. Northrop. J. Gen. Physiol. **24:** 213, 1940.
(108) Hesselvik, L. Ztschr. Physiol. Chem. **254:** 144, 1938.
(109) Hesselvik, L. Acta Med. Scand. **101:** 461, 1939.
(110) Hesselvik, L. Skand. Arch. Physiol. **82:** 151, 1939.
(111) Hesselvik, L. Acta Med. Scand. **105:** 153, 1940.
(112) Hewitt, L. F. Biochem. J. **30:** 2229, 1936; **31:** 360, 534, 1937; **32:** 26, 1938; **33:** 1496, 1939.
(113) Hoch, H. and C. J. O. R. Morris. Nature **156:** 234, 1945.
(114) Holmberg, C. G. and A. Grönwall. Hoppe-Seyler's Ztschr. **273:** 199, 1942.
(115) Hopkins, R. H., E. G. Stopher and D. E. Dolby. J. Inst. Brewing **46:** 426, 1940.
(116) Horsfall, F. L. Ann. N. Y. Acad. Sci. **34:** 203, 1939.
(117) Irving, G. W., Jr., T. D. Fontaine and R. C. Warner. Arch. Biochem. **7:** 475, 1945.
(118) Jacob, J. Experientia **2:** 110, 1946.
(119) Jameson, E. and C. Alvarez-Tostado. J. Phys. Chem. **43:** 1165, 1939.
(120) Jameson, E., C. Alvarez-Tostado and H. H. Sortor. Proc. Soc. Exper. Biol. and Med. **51:** 163, 1942.
(121) Jameson, E. and C. Alvarez-Tostado. J. Am. Chem. Soc. **65:** 459, 1943.
(122) Jorpes, E. and T. Thaning. Acta Physiol. Scand. **1:** 389, 1941.
(123) Kabat, E. A., D. H. Moore and H. Landow. J. Clin. Investigation **21:** 571, 1942.

(124) Kabat, E. A., F. M. Hanger, D. H. Moore and H. Landow. J. Clin. Investigation 22: 563, 1943.
(125) Kabat, E. A., H. Kaiser and H. Sikorski. J. Exper. Med. 80: 299, 1944.
(126) Kahler, H., M. J. Shear and J. L. Hartwell. J. Nat. Cancer Inst. 4: 123, 1943.
(127) Kass, E. H. Science 101: 337, 1945.
(128) Kekwick, R. A. and K. O. Pedersen. Biochem. J. 30: 2201, 1936.
(128a) Kekwick, R. A. Biochem. J. 32: 552, 1938.
(129) Kekwick, R. A. Biochem. J. 33: 1122, 1939.
(130) Kekwick, R. A. Biochem. J. 34: 1248, 1940.
(131) Kekwick, R. A. Trans. Faraday Soc. 36: 47, 1940.
(132) Kekwick, R. A. and B. R. Record. Brit. J. Exper. Pathol. 22: 29, 1940.
(133) Kekwick, R. A. Biochem. J. 37: 651, 1943.
(134) Keys, A., H. L. Taylor, O. Mickelson and A. Henschel. Science 103: 669, 1946.
(135) Kirkwood, J. G. J. Chem. Phys. 9: 878, 1941.
(136) Koenig, V. L. and K. R. Hogness. Arch. Biochem. 9: 119, 1946.
(137) Krebs, E. G. J. Lab. Clin. Med. 31: 85, 1946.
(138) Krejci, L. E., L. D. Smith and T. J. Dietz. J. Franklin Inst. 231: 396, 1941.
(139) Krejci, L. E., R. K. Jennings and L. D. Smith. J. Franklin Inst. 232: 592, 1941.
(140) Krejci, L. E. J. Franklin Inst. 234: 596, 1942.
(141) Krejci, L. E., R. K. Jennings and L. D. Smith. J. Immunol. 45: 111, 1942.
(142) Krejci, L. E., A. H. Stock, E. B. Sanigar and E. O. Kramer. J. Biol. Chem. 142: 785, 1942.
(143) Krejci, L. E., L. Sweeny and E. B. Sanigar. J. Biol. Chem. 158: 693, 1945.
(144) Lagercrantz, C. Uppsala Lakareforenings Forh. 51: 117, 1945.
(145) Landsteiner, K.; L. G. Longsworth and J. van der Scheer. Science 88: 83, 1938.
(146) Lee, C. D. and H. L. Wilcke. J. A. Vet. Med. Assoc. 94: 178, 1939.
(147) Lewis, J. H., C. S. Davidson, G. R. Minot, J. P. Soulier, H. J. Tagnon and F. H. L. Taylor. J. Clin. Investigation 25: 870, 1946.
(148) Lewis, L. A., E. P. McCullagh and J. Clark. Am. J. Med. Sci. 208: 727, 1944.
(149) Lewis, L. A., R. W. Schneider, E. P. McCullagh and J. Clark. J. Clin. Endocrinol. 4: 535, 1944.
(150) Lewis, L. A. J. Biol. Chem. 162: 473, 1946.
(151) Li, C. H., W. R. Lyons and H. M. Evans. Science 90: 622, 1939.
(152) Li, C. H., W. R. Lyons and H. M. Evans. J. Gen. Physiol. 23: 433, 1940.
(153) Li, C. H., H. M. Evans and D. H. Wonder. J. Gen. Physiol. 23: 733, 1940.
(154) Li, C. H., W. R. Lyons and H. M. Evans. J. Am. Chem. Soc. 62: 2925, 1940.
(155) Li, C. H. and H. L. Fraenkel-Conrat. J. Am. Chem. Soc. 64: 1586, 1942.
(156) Li, C. H., M. E. Simpson and H. M. Evans. J. Am. Chem. Soc. 64: 367, 1942.
(157) Li, C. H., H. M. Evans and M. E. Simpson. J. Biol. Chem. 149: 413, 1943.
(158) Li, C. H. J. Am. Chem. Soc. 66: 1795, 1944.
(159) Li, C. H., H. M. Evans and M. E. Simpson. J. Biol. Chem. 159: 353, 1945.
(160) Li, C. H. and W. O. Reinhardt. J. Biol. Chem. 167: 497, 1947.
(161) Linton, R. W., B. N. Mitra and S. C. Seal. Indian J. Med. Research 26: 329, 1938.
(162) Linton, R. W., L. D. Smith and L. E. Krejci. Arch. Biochem. 4: 195, 1944.
(163) Longsworth, L. G. Ann. N. Y. Acad. Sci. 34: 187, 1939.
(164) Longsworth, L. G. J. Am. Chem. Soc. 61: 529, 1939.
(165) Longsworth, L. G. and D. A. MacInnes. Chem. Rev. 24: 271, 1939.
(166) Longsworth, L. G., T. Shedlovsky and D. A. MacInnes. J. Exper. Med. 70: 399, 1939.
(168) Longsworth, L. G., R. K. Cannan and D. A. MacInnes. J. Am. Chem. Soc. 62: 2580, 1940.
(169) Longsworth, L. G. and D. A. MacInnes. J. Exper. Med. 71: 77, 1940.
(170) Longsworth, L. G. Chem. Rev. 30: 323, 1942.
(171) Longsworth, L. G. and D. A. MacInnes. J. Gen. Physiol. 25: 507, 1942.

(172) LONGSWORTH, L. G., R. M. CURTIS AND R. H. PEMBROKE, JR. J. Clin. Investigation 24: 46, 1945.
(173) LONGSWORTH, L. G. Ind. and Eng. Chem. Anal. Ed. 18: 219, 1946.
(174) LOURAU, M. J. Chim. Phys. 38: 149, 1941.
(175) LÖVGREN, O. Acta Med. Scand. Suppl. 163, 1945.
(176) LOZNER, E. L., S. LEMISH, A. S. CAMPBELL AND L. R. NEWHOUSER. Blood 1: 459. 1946.
(177) LUCK, J. M., C. C. NIMMO AND C. ALVAREZ-TOSTADO. Proc. Soc. Exper. Biol. and Med. 46: 151, 1941.
(178) LUETSCHER, J. A., JR. J. Am. Chem. Soc. 61: 2888, 1939.
(179) LUETSCHER, J. A., JR. J. Clin. Investigation 19: 313, 1940.
(180) LUETSCHER, J. A., JR. J. Clin. Investigation 20: 99, 1941.
(181) LUETSCHER, J. A., JR. J. Clin. Investigation 23: 365, 1944.
(182) LUETSCHER, J. A., JR. Unpublished observations.
(183) LUNDGREN, H. P., D. W. ELAM AND R. A. O'CONNELL. J. Biol. Chem. 149: 183, 1943.
(184) MACHEBOEUF, M. A. Compt. rend. soc. biol. 135: 1241, 1941.
(185) MACHEBOEUF, M. A. AND H. VANAUD. Compt. rend. soc. biol. 135: 1249, 1941.
(186) MACINNES, D. A. AND L. G. LONGSWORTH. Colloid chemistry, theory and methods. Reinhold Publ. Corp., New York, 5: 387, 1944.
(187) MACPHERSON, C. F. C., D. H. MOORE AND L. G. LONGSWORTH. J. Biol. Chem. 156: 381, 1944.
(188) MACPHERSON, C. F. C., M. HEIDELBERGER AND D. H. MOORE. J. Am. Chem. Soc. 67: 578, 1945.
(189) MADDEN, S. C. AND G. H. WHIPPLE. Physiol. Rev. 20: 194, 1940.
(190) MALMROS, H. AND G. BLIX. Acta Med. Scand. Suppl. 170, 1946.
(191) MARTIN, N. H. Brit. J. Exper. Path. 27: 363, 1946.
(192) MASUCCI, P. AND R. J. DE FALCO. Proc. Soc. Exper. Biol. and Med. 58: 67, 1945.
(193) MATUURA, Y. J. Oriental. Med. 31: 715, 1939.
(194) MCCULLAGH, E. P., L. A. LEWIS AND W. F. OWEN. Cleveland Clin. Quart. 10: 88, 1943.
(195) MCCULLAGH, E. P., L. A. LEWIS AND J. CLARK. Am. J. Med. Sci. 210: 81, 1945.
(196) MCFARLANE, A. S. Biochem. J. 29: 1209, 1935.
(197) MCFARLANE, A. S. AND R. A. KEKWICK. Biochem. J. 32: 1607, 1938.
(198) MCFARLANE, A. S. Trans. Faraday Soc. 36: 257, 1940.
(199) MCMEEKIN, T. L. J. Am. Chem. Soc. 62: 3393, 1940.
(200) MELLANDER, O. Biochem. Ztschr. 300: 240, 1939.
(201) MELLANDER, O. Nature 155: 604, 1945.
(202) MELNICK, J. L. AND K. G. STERN. Enzymologia 8: 129, 1940.
(203) MENKIN, V. Science 101: 422, 1945.
(204) MILLER, G. L., M. A. LAUFFER AND W. M. STANLEY. J. Exper. Med. 80: 549, 1944.
(205) MIRKSY, A. E. AND A. W. POLLISTER. Proc. Natl. Acad. Sci. U. S. 28: 344, 1942.
(206) MODERN, F., O. REPETTO AND G. RUFF. Rev. soc. Argentina biol. 16: 544, 1941.
(207) MOONEY, M. Temperature. Its measurement and control in science and industry. Reinhold Publishing Co., New York, p. 428, 1941.
(208) MOORE, D. B., P. S. PIERSON, F. M. HANGER AND D. H. MOORE. J. Clin. Investifation 24: 292, 1945.
(209) MOORE, D. H., J. VAN DER SCHEER AND R. W. G. WYCKOFF. J. Immunol. 38: 221, 1940.
(210) MOORE, D. H. AND J. LYNN. J. Biol. Chem. 141: 819, 1941.
(211) MOORE, D. H. J. Am. Chem. Soc. 64: 1090, 1942.
(212) MOORE, D. H., E. A. KABAT AND A. B. GUTMAN. J. Clin. Investigation 22: 67, 1943.
(213) MOORE, D. H. Optical methods in electrophoretic and ultracentrifugal analyses. Medical Physics. Year Book Publishers, Chicago, 1944.
(215) MOORE, D. H., L. LEVIN AND J. H. LEATHEM. J. Biol. Chem. 153: 349, 1944.
(216) MOORE, D. H. AND M. MAYER. J. Biol. Chem. 156: 777, 1944.

(217) MOORE, D. H. AND L. REINER. J. Biol. Chem. 156: 411, 1944.
(218) MOORE, D. H. J. Biol. Chem. 161: 21, 1945.
(219) MOORE, D. H., L. LEVIN AND G. K. SMELSER. J. Biol. Chem. 157: 723, 1945.
(220) MOORE, D. H., S. C. SHEN AND C. S. ALEXANDER. Proc. Soc. Exper. Biol. and Med. 58: 307, 1945.
(221) MOYER, L. S. J. Phys. Chem. 42: 71, 1938.
(222) MOYER, L. S. AND M. H. GORIN. J. Biol. Chem. 133: 605, 1940.
(223) MOYER, L. S. AND E. Z. MOYER. J. Biol. Chem. 132: 373, 1940.
(224) MUNRO, F. L. AND M. P. MUNRO. J. Clin. Investigation 25: 814, 1946.
(225) MUNRO, M. P. AND F. L. MUNRO. J. Biol. Chem. 150: 427, 1943.
(226) MUNRO, M. P. AND J. E. THOMAS. Am. J. Physiol. 145: 140, 1945.
(227) MUNRO, M. P. AND A. AVERY. Am. J. Physiol. 146: 673, 1946.
(228) MUSSIO FOURNIER, J. C., O. CONTI AND J. C. LABORDE. Proc. Soc. Exper. Biol. and Med. 45: 493, 1940.
(229) NEURATH, H., G. R. COOPER, D. G. SHARP, A. R. TAYLOR, D. BEARD AND J. W. BEARD. J. Biol. Chem. 140: 293, 1941.
(230) NEURATH, H. AND F. W. PUTNAM. J. Biol. Chem. 160: 397, 1945.
(231) NEURATH, H., E. VOLKIN, J. O. ERICKSON, F. W. PUTNAM, H. W. CRAIG, G. R. COOPER, D. G. SHARP, A. R. TAYLOR AND J. W. BEARD. Science 101: 68, 1945.
(232) NEWELL, J. M., A. STERLING, M. F. OXMAN, S. S. BURDEN AND L. E. KREJCI. J. Allergy 10: 513, 1939.
(233) NEWELL, J. M. J. Allergy 14: 444, 1943.
(234) NIELSEN, L. E. AND J. G. KIRKWOOD. J. Am. Chem. Soc. 68: 181, 1946.
(235) OLHAGEN, B. Studies on thermostable anticomplementary human sera. Dissertation. Stockholm, 1945.
(236) ORLOV, G. A. Z. Microbiol. Epidemiol. Immunitatsforsch. (U.S.S.R.) 19: 86, 1937.
(237) ORR, W. F., JR. AND D. H. MOORE. Proc. Soc. Exper. Biol. and Med. 46: 357, 1941.
(238) PACKCHANIAN, A. Texas Repts. Biol. Med. 1: 157, 1943.
(239) PAPPENHEIMER, A. M., H. P. LUNDGREN AND J. W. WILLIAMS. J. Exper. Med. 71: 247, 1940.
(240) PASYNSKII, A. AND I. PETROV. J. Phys. Chem. (U.S.S.R.) 12: 566, 1938.
(241) PEDERSEN, K. O. Nature 154: 575, 1944.
(242) PEDERSEN, K. O. Ultracentrifugal studies on serum and serum fractions. Dissertation. Almqvist and Wiksells, Uppsala, 1945.
(243) PERLMAN, E., J. G. BULLOWA AND R. GOODKIND. J. Exper. Med. 77: 97, 1943.
(244) PERLMANN, G. E., W. W. L. GLENN AND D. KAUFMAN. J. Clin. Investigation 22: 627, 1943.
(245) PERLMANN, G. E. AND D. F. KAUFMAN. J. Am. Chem. Soc. 67: 638, 1945.
(246) PERLMANN, G. E. AND D. F. KAUFMAN. J. Clin. Investigation 25: 931P. 1946.
(247) PETERMANN, M. L. AND N. V. HAKALA. J. Biol. Chem. 145: 701, 1942.
(248) PETERMANN, M. L. J. Am. Chem. Soc. 68: 106, 1946.
(249) PETERMANN, M. L. J. Biol. Chem. 162: 37, 1946.
(250) PHILPOT, J. ST. L. Nature 141: 283, 1938.
(251) PHILPOT, J. ST. L. Trans. Faraday Soc. 36: 38, 1940.
(252) PILLEMER, L., E. E. ECKER, J. L. ONCLEY AND E. J. COHN. J. Exper. Med. 74: 297, 1941..
(253) PILLEMER, L., J. L. ONCLEY, M. MELIN, J. ELLIOTT AND M. C. HUTCHINSON. J. Clin. Investigation 23: 550, 1944.
(254) PILLEMER, L. AND M. C. HUTCHINSON. J. Biol. Chem. 158: 299, 1945.
(255) PLENTL, A. A. AND I. PAGE. J. Biol. Chem. 147: 143, 1943.
(256) POLSON, A. Nature 145: 27, 1940.
(257) POLSON, A. Onderstepoort J. Vet. Sci. Animal Ind. 16: 51, 1941.
(258) POLSON, A. Onderstepoort J. Vet. Sci. Animal Ind. 20: 159, 1945.
(259) POWNEY, J. AND L. J. WOOD. Trans. Faraday Soc. 36: 420, 1940.

(260) PUTNAM, F., J. O. ERICKSON, E. VOLKIN AND H. NEURATH. J. Gen. Physiol. 26: 513, 1943.
(261) PUTNAM, F. AND H. NEURATH. J. Biol. Chem. 159: 195, 1945.
(262) PUTNAM, F. W. AND H. NEURATH. J. Biol. Chem. 160: 239, 1945.
(263) PUTZEYS, P. AND P. VAN DE WALLE. Trans. Faraday Soc. 36: 32, 1940.
(264) QUENSEL, O. Untersuchungen uber die Gerstenglobuline. Dissertation. Almqvist and Wiksell, Uppsala, 1942.
(265) REINER, L., D. H. MOORE, E. H. LONG AND M. GREEN. J. Biol. Chem. 146: 583, 1942.
(267) ROSS, V., D. H. MOORE AND E. G. MILLER, JR. J. Biol. Chem. 144: 667, 1942.
(268) ROSS, V., E. G. MILLER, JR., D. H. MOORE AND H. SIKORSKI. Proc. Soc. Exper. Biol. and Med. 54: 179, 1943.
(269) ROTHEN, A. J. Gen. Physiol. 24: 203, 1940.
(270) ROTHEN, A. J. Gen. Physiol. 25: 487, 1942.
(271) ROY, D. K. Ann. Biochem. Exper. Med. 3: 39, 1943.
(272) RUTSTEIN, D. D., F. H. CLARKE AND L. M. TARAN. Science 101: 669, 1945.
(273) SAMEK, M., C. NUCIC AND V. PIRKMAIER. Kolloid Ztschr. 94: 350, 1941.
(274) SAN CLEMENTE, C. L. AND I. F. HUDDLESON. Mich. State College Agr. Exper. Sta., Tech. Bull. 182: 3, 1943.
(275) SANDERS, E., I. F. HUDDLESON AND P. J. SCHAIBLE. J. Biol. Chem. 155: 469, 1944.
(276) SANIGAR, E. B., L. E. KREJCI AND E. O. KRAEMER. J. Franklin Inst. 235: 298, 1943.
(277) SAYERS, G., A. WHITE AND C. N. H. LONG. J. Biol. Chem. 149: 425, 1943.
(278) SCHNEIDER, R. W., L. A. LEWIS AND E. P. MCCULLAGH. Am. J. Med. Sci. 212: 462, 1946.
(279) SCUDDER, J. Ann. Surg. 112: 502, 1940.
(280) SCHWERT, G. W., F. W. PUTNAM AND D. R. BRIGGS. Arch. Biochem. 4: 371, 1944.
(281) SEEGERS, W. H., E. C. LOOMIS AND J. M. VANDERBILT. Proc. Soc. Exper. Biol. and Med. 56: 70, 1944.
(282) SEIBERT, F. B., K. O. PEDERSEN AND A. TISELIUS. J. Exper. Med. 68: 413, 1938.
(283) SEIBERT, F. B. J. Biol. Chem. 133: 593, 1940.
(284) SEIBERT, F. B. AND D. W. WATSON. J. Biol. Chem. 140: 55, 1941.
(285) SEIBERT, F. B. AND J. T. GLENN. Am. Rev. Tuberculosis 44: 9, 1941.
(286) SEIBERT, F. B. AND J. W. NELSON. Proc. Soc. Exper. Biol. and Med. 49: 77, 1942.
(287) SEIBERT, F. B. AND J. W. NELSON. J. Biol. Chem. 143: 29, 1942.
(288) SEIBERT, F. B. AND J. W. NELSON. Am. Rev. Tuberculosis 47: 66, 1943.
(289) SEIBERT, F. B. AND J. W. NELSON. J. Am. Chem. Soc. 65: 272, 1943.
(290) SEIBERT, F. B., M. V. SEIBERT, A. J. ATNO AND H. W. CAMPBELL. J. Clin. Investigation 26: 90, 1947.
(291) SENEKJIE, H. A. AND L. C. SCOTT. Proc. Soc. Exper. Biol. and Med. 51: 174, 1942.
(292) SHAPIRO, S., V. ROSS AND D. H. MOORE. J. Clin. Investigation 22: 137, 1943.
(293) SHARP, D. G., A. R. TAYLOR, D. W. BEARD AND J. W. BEARD. J. Biol. Chem. 142: 193, 1942.
(294) SHARP, D. G., G. R. COOPER AND H. NEURATH. J. Biol. Chem. 142: 203, 1942.
(295) SHARP, D. G., M. H. HEBB, A. R. TAYLOR AND J. W. BEARD. J. Biol. Chem. 142: 217, 1942.
(296) SHARP, D. G., G. R. COOPER, J. O. ERICKSON AND H. NEURATH. J. Biol. Chem. 144: 139, 1942.
(297) SHARP, D. G., A. R. TAYLOR, D. W. BEARD AND J. W. BEARD. J. Immunol. 44: 115, 1942.
(298) SHARP, D. G., A. R. TAYLOR, D. W. BEARD AND J. W. BEARD. Proc. Soc. Exper. Biol. and Med. 50: 358, 1942.
(299) SHEDLOVSKY, T., A. ROTHEN, R. O. GREEP, H. B. VAN DYKE AND B. F. CHOW. Science 92: 178, 1940.
(300) SHEDLOVSKY, T. AND J. E. SMADEL. J. Exper. Med. 72: 511, 1940.
(301) SHEDLOVSKY, T. AND J. SCUDDER. J. Exper. Med. 75: 119, 1942.

(302) Shedlovsky, T. Ann. N. Y. Acad. Sci. **43**: 259, 1943.
(303) Shemin, D., E. E. Sproul and J. W. Jobling. J. Exper. Med. **72**: 697, 1940.
(304) Shipley, R. A., K. G. Stern and A. White. J. Exper. Med. **69**: 785, 1939.
(305) Shreder, V. N. Biol. Zhur. **5**: 657, 1936.
(306) Shreder, V. N. Bull. acad. sci. U.R.S.S., Ser. biol. 426, 1940.
(307) Smadel, J. E., E. G. Pickels, T. Shedlovsky and T. M. Rivers. J. Exper. Med. **72**: 523, 1940.
(308) Smadel, J. E. and T. Shedlovsky. Ann. N. Y. Acad. Sci. **43**: 35, 1942.
(309) Smetana, H. and D. Shemin. J. Exper. Med. **73**: 223, 1941.
(310) Smith, E. C. B. Biochem. J. **34**: 1122, 1940.
(311) Smith, E. C. B. Biochem. J. **34**: 1176, 1940.
(312) Smith, E. L. J. Biol. Chem. **164**: 345, 1946.
(313) Smith, E. L. J. Biol. Chem. **165**: 665, 1946.
(314) Smith, O. W. and G. U. S. Smith. Science **102**: 253, 1945.
(315) Sookne, A. M. and M. Harris. Textile Research **9**: 374, 437, 1939.
(316) Spies, J. R. J. Am. Chem. Soc. **63**: 1166, 1941.
(317) Spies, J. R., H. S. Bernton and H. Stevens. J. Am. Chem. Soc. **63**: 2163, 1941.
(318) Stats, D., E. Perlman, J. G. M. Bullowa and R. Goodkind. Proc. Soc. Exper. Med. **53**: 188, 1943.
(319) Stearns, T. W. and M. H. Roepke. J. Bact. **42**: 411, 1941.
(320) Stenhagen, E. Biochem J. **32**: 714, 1938.
(321) Stenhagen, E. and T. Teorell. Nature **141**: 415, 1938.
(322) Stenhagen, E. and T. Teorell. Trans. Faraday Soc. **35**: 743, 1939.
(323) Stern, K. G. Ann. N. Y. Acad. Sci. **34**: 147, 1939.
(324) Stern, K. G. Biol. Symposia **10**: 291, 1943.
(325) Stern, K. G. J. Biol. Chem. **152**: 345, 1944.
(326) Stern, K. G., M. Reiner and R. H. Silber. J. Biol. Chem. **161**: 731, 1945.
(327) Stern, K. G., A. H. Schein and J. S. Wallerstein. J. Biol. Chem. **166**: 59, 1946.
(328) Stern, K. G. and M. Reiner. Yale J. Biol. Med. **19**: 67, 1946.
(329) Svedberg, T. and K. O. Pedersen. The ultracentrifuge. Oxford University Press, London, 1940.
(330) Svensson, H. Kolloid Ztschr. **87**: 181, 1939.
(331) Svensson, H. Kolloid Ztschr. **90**: 141, 1940.
(332) Svensson, H. J. Biol. Chem. **139**: 805, 1941.
(333) Svensson, H. Arkiv f. Kemi, mineral., o. geol. **15B**: no. 19, 1942.
(334) Svensson, H. Chem. Zentr. **2**: 76, 1942.
(335) Svensson, H. Arkiv f. Kemi, mineral., o. geol. **17A**: no. 5, 1943.
(336) Svensson, H. Arkiv f. Kemi, mineral., o. geol. **17A**: no. 14, 1943.
(337) Svensson, H. The Svedberg Mem. Vol. 1944, 213.
(338) Svensson, H. Arkiv f. Kemi, mineral., o. geol. **22A**: no. 10, 1946.
(339) Swartz, N. Nord. Med. **18**: 673, 1943; **23**: 1713, 1944.
(340) Taylor, A. R., D. G. Sharp, D. Beard and J. W. Beard. Proc. Soc. Exper. Biol. and Med. **51**: 137, 1942.
(341) Taylor, H. L. and A. Keys. J. Biol. Chem. **148**: 379, 1943.
(342) Theorell, H. Biochem. Ztschr. **275**: 1, 1934.
(343) Theorell, H. Biochem. Ztschr. **278**: 291, 1935.
(344) Theorell, H. and A. Akesson. J. Am. Chem. Soc. **63**: 1804, 1941.
(345) Thompson, K. W. and J. L. Melnick. Endocrinology **28**: 723, 1941.
(346) Thorn, G. W., S. H. Armstrong, Jr., V. D. Davenport, L. M. Woodruff and F. H. Tyler. J. Clin. Investigation **24**: 802, 1945.
(347) Thorn, G. W., S. H. Armstrong, Jr. and V. D. Davenport. J. Clin. Investigation **25**: 304, 1946.
(348) Tiselius, A. Trans. Faraday Soc. **33**: 524, 1937.
(349) Tiselius, A. Biochem. J. **31**: 313, 1937.

(350) Tiselius, A. Biochem J. **31**: 1464, 1937.
(351) Tiselius, A. Kolloid Ztschr. **85**: 129, 1938.
(352) Tiselius, A., G. E. Henschen and H. Svensson. Biochem. J. **32**: 1814, 1938.
(353) Tiselius, A. and E. A. Kabat. Science **87**: 416, 1938.
(354) Tiselius, A. and F. L. Horsfall, Jr. Arkiv f. Kemi, mineral o. geol. **13A**: no. 18, 1939.
(355) Tiselius, A. Rept. Proc. 3rd intern. Congr. Microbiol. 1939, 54, 1940.
(356) Tiselius, A. and I. B. Eriksson-Quensel. Biochem. J **33**: 1752, 1939.
(357) Tiselius, A. and F. L. Horsfall, Jr. J. Exper. Med. **69**: 83, 1939.
(358) Tiselius, A. and E. A. Kabat. J. Exper. Med. **69**: 119, 1939.
(359) Tiselius, A. Harvey Lectures **35**: 37, 1940.
(360) Tiselius, A. and H. Svensson. Trans. Faraday Soc. **36**: 16, 1940.
(361) Tiselius, A. Svensk. Kem. Tid. **53**: 305, 1941.
(362) Tiselius, A. and O. Dahl. Arkiv. Kemi. Mineral. o. geol. **14B**: no. 31, 1941.
(363) Tiselius, A. and A. Gronwall. Arkiv. f. Kemi, Mineral., o. geol. **17A**: no. 13, 1943.
(364) Treffers, H. P., D. H. Moore and M. Heidelberger. J. Exper. Med. **75**: 135, 1942.
(365) Van der Scheer, J. and R. W. G. Wyckoff. Science **91**: 485, 1940.
(366) Van der Scheer, J., R. W. G. Wyckoff and F. H. Clarke. J. Immunol. **39**: 65, 1940.
(367) Van der Scheer, J., R. W. G. Wyckoff and F. H. Clarke. J. Immunol. **40**: 39, 1941.
(368) Van der Scheer, J., R. W. G. Wyckoff and F. H. Clarke. J. Immunol. **40**: 173, 1941.
(369) Van der Scheer, J., J. B. Lagsdin and R. W. G. Wyckoff. J. Immunol. **41**: 209, 1941.
(370) Van der Scheer, J., R. W. G. Wyckoff and F. H. Clarke. J. Immunol. **41**: 349, 1941.
(371) Van der Scheer, J., E. Bohnel, F. H. Clarke and R. W. G. Wyckoff. J. Immunol. **44**: 165, 1942.
(372) Verway, W. F. and M. Frobisher, Jr. Am. J. Hyg. **32B**: 55, 63, 1940.
(373) Wainwright, G. W. and B. D. Davis. Unpublished observations.
(374) Weibull, C. and A. Tiselius. Arkiv. f. Kemi., Mineral., geol. **19A**: no. 19, 1945.
(375) West, R. and D. H. Moore. Trans. Assoc. Am. Physicians **57**: 259, 1942.
(376) White, A., R. W. Bonsnes and C. N. H. Long. J. Biol. Chem. **143**: 447, 1942.
(377) White, A. Publ. A. A. A. S., pg. 1, 1944.
(378) White, A. and T. F. Dougherty. Proc. Soc. Exper. Biol. and Med. **56**: 26, 1944.
(379) White, A. and T. F. Dougherty. Endocrinol. **36**: 207, 1945.
(380) Wilander. Skand. Arch. Physiol., Suppl. 81.
(381) Williams, J. W., M. L. Petermann, G. C. Colvas, M. B. Goodloe, J. L. Oncley and S. H. Armstrong, Jr. J. Clin. Investigation **23**: 433, 1944.
(382) Wright, G. G. and J. Oliver-Gonzalez. J. Infect. Dis. **72**: 242, 1943.
(383) Wyckoff, R. W. G. and M. Rhian. J. Immunol. **51**: 359, 1945.
(384) Zeldis, L. J. and E. L. Alling. J. Exper. Med. **81**: 515, 1945.
(385) Zeldis, L. J., E. L. Alling, A. B. McCoord and J. P. Kulka. J. Exper. Med. **82**: 157, 1945.
(386) Zeldis, L. J., E. L. Alling, A. B. McCoord and J. P. Kulka. J. Exper. Med. **82**: 411, 1945.
(387) Ziff, M. and D. H. Moore. J. Biol. Chem. **153**: 653, 1944.
(388) Zittle, C. A. and F. B. Seibert. J. Immunol. **43**: 47, 1942.

PROBLEMS IN INVERTEBRATE ELECTROPHYSIOLOGY

THEODORE HOLMES BULLOCK

Department of Zoology, University of California, Los Angeles

Many aspects of the electrophysiological study of invertebrate nervous systems have been covered in recent reviews. Thus, the useful articles of Prosser (121) and Welsh and Schallek (149) specifically deal with invertebrates while those of Bishop (16) and Grundfest (64) include much recent work on these forms. Conduction and excitability in nerve are well treated by Curtis and Cole (46, see also Hodgkin and Rushton, 71). Wiersma (153) has dealt with muscle physiology in invertebrates. The rôle of acetylcholine is discussed in Nachmansohn (100) and Gerard (56). Chemical excitation of peripheral nerve is reviewed by Brink, Bronk and Larrabee (25).

Recent advances in certain areas of investigation have not, however, come to general attention. The present paper undertakes to summarize the present state of knowledge in four of the important fields of application of electrophysiological methods: the comparative study of synaptic mechanisms, physiological neuronography—which has been applied especially to giant nerve fiber systems, the phenomenon of spontaneous central activity and the identification and analysis of sensory structures. In each case particular attention is directed to problems and opportunities for further investigation.

Junctional transmission. As a result of recent work, the subject of comparative synaptology promises particularly interesting contributions to general neurology. Whereas the advantages of lower forms for the study of conduction lie in the apparent uniformity of the mechanism, much of the advantage for synaptology lies in the diversity of the manifestation of a perhaps common fundamental physico-chemical process. The diversity appears to rest on quantitative differences in physiologic properties and on differences in anatomical arrangement at the junction.

Nerve nets. Although electrical recording technics have not yet been successfully applied to nerve nets, it appears appropriate to discuss briefly the status of our understanding of this important type of nervous organization. The most primitive nervous systems known, those of the coelenterates, have for long been regarded as asynaptic. In order to explain local responses or subtotal response of any kind it was necessary to assume decremental conduction (Parker, 105), a phenomenon for which no independent evidence was cited and which was never successfully applied to the explanation of diversified behavior patterns in higher coelenterates. In 1935, as a result of the first modern physiological analysis of these systems, Pantin (103) offered an alternative explanation depending on the existence of discontinuities in the nerve net which require facilitation by successive impulses at critical intervals. Such neuro-neural junctions appear to be absent or permanently facilitated in extensive regions of the net, the specific reactions to stimuli depending on differentiated properties of neuro-myal junc-

tions in different muscles. But neuro-neural junctions were found, physiologically, by the dependence on facilitation. Junctions which depend greatly on facilitation are probably widespread among invertebrates. Though good evidence is at hand only for anthozoa and crustacea (Pantin, 104; Wiersma, 153) suggestive evidence has led Pantin to include echinoderms and molluscs (104) and Botsford to add annelids (21). Pantin's picture accords with available facts and explains observed responses in a satisfactory way (Pantin, 104; Bullock, 29). Anatomical evidence is notoriously difficult to obtain and to rely on in this matter but the reports of discontinuity in the net (Bozler, 23; Woolard and Harpman, 158) are generally given greater weight today than those (96, 105) of continuity (66, 74, 33, 104, 121, 42).

The importance of the question of the existence of synapses in the nerve net rests on the fact that if they are there, the well-known diffuse conduction in the net forces the acceptance of unpolarized synapses, unless secondary assumptions of duplicate paths in every direction are to be made. The balance of evidence today, although certainly not ruling out continuity in certain parts of the nerve net, indicates that unpolarized synapses exist in the nerve nets of coelenterates and probably enteropneusts and other groups (Bullock, 27, 28, 33). Nerve net in this usage becomes a physiological term denoting a diffuse locally autonomous plexus conducting in all directions. In the present state of knowledge the entities classically called nerve nets must be characterized in these terms since the exact anatomical status is unknown. In any case, loose usage of the term to include simple plexuses of peripheral terminals of neurons whether the cyton is far away as in human skin (15, 91) or in the periphery as in insects, worms and possibly other forms (155) must be deplored, even if anastomoses between sensory fibers are found. Unless evidence is available that the plexus can mediate responses locally, autonomously, is thus a distributing as well as receiving path, and can conduct diffusely, it should not be called a nerve net.

It may be hoped that electrical recording technics will be applied to these nervous systems. When they are, crucial contributions to the question of their functional organization and to the general understanding of common denominators versus derived properties of nervous tissue may be expected.

Even more neglected by modern physiological technics are the synapses and neuro-myal junctions in the several phyla of animals possessing ganglionated nervous systems simpler in organization than those of articulates and molluscs. Considering the great advances in the evolution of nervous function between the lowest and the highest invertebrates it would seem that efforts to analyze the properties of nervous units in these intermediate forms would be worth while.

"Typical" synapses in arthropods. Until recent months the only definite central synapses in invertebrates which had received careful study by the technics of electrical recording were those in the crayfish preparation of Prosser (111, 114, 119) and the cockroach preparation of Pumphrey and Rawdon-Smith (125).

Prosser's preparation includes the first synapses in the proprioceptive and tactile pathways, in particular those between primary sensory cells in the tactile hairs of the tail and second order neurons in the last abdominal ganglion. A

special advantage of this situation is the possibility of stimulating separately each of many single presynaptic units, permitting quantitative study of spatial summation, which in these junctions, in contrast to several others among the invertebrates, is necessary for transmission. Although the pathways are relatively fast, synaptic delay is long (2–8ms.). Transmission cannot follow above ten per second but facilitation is exhibited. After-discharge of the post-synaptic units is not marked. The preparation lends itself to the study of effects of external agents and a variety of drugs and salts has been tested (119). The intimate anatomy is still unknown.

The synapse in the last abdominal ganglion of the cockroach between primary vibration receptor afferents from the cercal nerve and large ascending second order neurons (125) presents an interesting contrast. Spatial summation is believed not to be necessary for the lowest threshold synapses in the rested state, but for all others (there are 12–16 second order fibers) supraminimal presynaptic volleys are necessary. Delay is as short as that in any natural junction known, including central synapses of mammals (but excluding the giant fiber septa of annelids and crustacea, whose status is still uncertain; see below), i.e., 0.6 ms. (Roeder et al., 133) The delay is much longer with weaker afferent volleys. Transmission can follow up to 50/sec. at high intensities of presynaptic stimulation, lower frequencies under submaximal stimulation, without showing any effects of the successive impulses. At higher frequencies some post-fibers drop out and delay is increased, although the remaining fibers continue to transmit in a one-to-one manner. In this semi-adapted condition facilitation occurs, as shown by recruitment of units upon shortening the interval between stimuli. After-discharge can be elicited (133). Lowenstein (93) has used time of abolition of response to air-puff stimulation in a method of bioassay of pyrethrum extracts, but it has not been established where, between receptor and second order axon, the compound acts. Roeder et al. (133) have used this preparation to study the effects of DFP, apparently the first test of this important anticholinesterase on single synapses. He reports a high sensitivity with increase in excitability and after-discharge as the first signs of effect. Later the synapses alternate between periods of complete block and periods of hyperactivity. Recovery from the DFP effect was not complete under the conditions used, but experience with this compound in other cases suggests that more acute experiments, using higher concentrations for shorter periods, may produce more marked effects with complete reversibility (40).

The properties of these synapses correspond in general to the usual textbook description and there is no reason to doubt at present that they are representative of most synapses in invertebrates. There exist, however, certain exceptional neurons in some of the higher invertebrates which display junctions of such character that their status as synapses is even uncertain. Special interest attaches to these cases and we may therefore turn to a brief review of the problem.

Annelid giant fiber discontinuities. Stough in 1926 (143) described and illustrated conspicuous septa across each of the three well-known giant nerve fibers in the dorsum of the earthworm nerve cord. The septa were shown to be com-

plete, oblique, simple partitions at segmental intervals. The unit of giant fiber between each two septa has one or several appended cell bodies so that the giant system can be thought of as a chain of compound neurons. Stough applied the term macrosynapse to these large, simple geometrically unpolarized junctions. Eccles, Granit and Young (49) showed by electrical recording that they conduct equally well in both directions and at a speed which permits essentially no time to be ascribed to delay at each junction. This has been confirmed (Rushton, 139, 136, 138; Bullock, 31, 32). As yet no physiologic sign of the septa has been reported although it is to be anticipated that such conspicuous discontinuity and differentiation of the fibers will be found to leave some mark in the conduction properties.

The question whether they must be regarded as synapses depends in the first place on the reality of the physiologic membrane and in the second upon the definition of synapse. Experimental demonstration of a functional membrane has not been made. But the considerable thickness of the histological membrane (143), the differentiation and interruption of neurofibrils (22, 142, 102, 32), and the presence of a clearly distinguishable doubly refractile layer continuous with the myelin around the fiber (144) suggest that the barrier is real. If this proves to be true the usual definition of a synapse as a region of functional contact between two discrete neuronal units will apply and the properties of polarization and delay common to most synapses could not be regarded as defining qualities. Instead they might be regarded as the results of a special anatomical arrangement (McCulloch, 97). The question may be met by adopting a more restrictive definition of synapse as is done by Bishop (16) who implies that a functional connection between discrete units cannot be regarded as a synapse until it has been shown to act as a switch which under natural conditions is not always closed. The utility of this criterion in specific cases may be questioned.

Reasons for doubting the reality of the membranes at the septa have been presented by Schoepfle (141). The argument is that the potential detected in the conducting medium around a nerve or muscle as a propagated wave of surface negativity approaches, is initially positive and immediately beyond the end of such a linear element is not followed by any negativity. Hence only an anelectrotonic or depressing effect could be exerted on another element whose excitable part was confined to this region of the external field of the prefiber, i.e., end-to-end junctions cannot be expected to transmit excitation but only inhibition. The argument is cogent if the geometry of the actual case corresponds to that of the theoretical one treated. A rigid analysis of the volume conductor problems involved with special reference to the expected effects of the oblique contacts (overlapping and tapering), in this case and other more complex forms found in other natural junctions is much needed for the proper evaluation of Schoepfle's argument and for the extension to actual situations of the schemes upon which Eccles has recently based a theory of synaptic transmission.

Eccles' theory (47, 48) also throws doubt on the reality of the septa of the earthworm as physiologic cell membranes. According to this concept a significant delay is theoretically necessary at any real junction, i.e., one where a mem-

brane like that around the axon interrupts the path of conduction. Apparently not all neurophysiologists are ready to accept that this is demanded by the membrane theory for all possible junctions (personal communications), but in any case it makes more important the precise determination of the degree of physiological discontinuity represented by the septa of the earthworm giant fibers. Conceivably they may provide a crucial test of Eccles' statement. The significance in this connection of the extremely short delays, often 0.1 ms., at the artificial synapses formed by the cut end of nerves or central tracts (Barron, 12; Renshaw and Therman, 127; Granit and Skoglund, 61), appears not to have been discussed. They are not readily attributed to subterminal interaction, which would have to occur 10 to 20 mm. from the cut surface in order to allow 0.5 ms. for delay, but neither has such a site been conclusively eliminated. Proper experiments for a clear localization of the synapse would seem to be practicable.

The septal system in the earthworm is not an isolated case. Among polychaetes giant fibers with cell bodies distributed along the length of the cord are common (90, 65, 54, 10). Septa have been described in only one form (143) but there have been no studies in which such reports might be expected since the first discovery of septa by Stough. It seems likely that many of the giant systems with distributed cell bodies will prove to have septa. One case is of special interest since the available anatomical information describes a typical synapse in each segment, with branching terminations of one fiber ending on the main stem of another (medial giant fiber pair of *Nereis*, Hamaker, 65). Yet conduction in these fibers and in all others tested in a recent study (15 species) is unpolarized and delays can be at most very short (35, 38).

The lateral giant fibers of the crayfish (85) and prawn (72) have definite septa and conduct both ways (Wiersma, 152; Prosser, 121). According to Wiersma (154) not more than 0.1 ms. can be attributed to delay at each.

A special problem is presented by the junctions between these lateral giant fibers. Both in crayfish (Wiersma, 154) and earthworm (Rushton, 136) a considerable delay (0.8 ms.) is introduced when an impulse crosses from one lateral to the other. There appears no reason to suppose the junction is polarized but it is easily fatigued in the crayfish. Rushton finds that in the earthworm impulses in the laterals which are thus made out of phase by 0.8 ms. somehow get into step if a sufficient length of fiber is available. Can electrotonic interaction of the type of Katz and Schmitt (87) be invoked—slowing down the leading or speeding up the trailing impulse?

Artificial synapses. Recently increased significance has been placed on studies of so-called artificial synapses. On the one hand they have been the only preparations available in which the properties of single junctions could be readily studied until Kuffler obtained isolated nerve muscle junctions (88). On the other hand they offer the valuable opportunity of varying the geometrical relations at the junction. These advantages have been best exploited in the work of Arvanitaki (8) on "ephapses" of squid giant fibers. She demonstrated that the pre-fiber could leave the post-fiber in an excited state only when it was arranged spatially in such a way that the impulse was not conducted past the

junction, thus permitting anelectrotonic depression to terminate the action. When the impulse in the pre-fiber stops soon after reaching the junction its final action is catelectrotonic and excitation takes place. The behavior of the preparations argues strongly for a purely electrical transmitter action and accounts quantitatively for the observed delay. There would appear to be no reason why an appropriate geometric arrangement of the elements could not provide an ephapse conducting in both directions, although this point was not tested. The properties of this preparation were particularly important in the elaboration of Eccles' (47, 48) picture of synaptic transmission.

Other artificial synapses which have been studied include the juxtaposed fibers of crab nerve (Arvanitaki, 7; Jasper and Monnier, 84; Katz and Schmitt, 87), and vertebrate nerve prepared in various ways to measure the effects of active fibers upon inactive ones (Marazzi and Lorente de No, 95; Bishop, 14). A preparation of special interest is the synapse formed at the cut end of nerve where impulses in fibers from the ventral root may be transmitted to fibers of the dorsal root (Granit and Skoglund, 61) to a degree depending on a complex of factors and with an apparent delay as short as 0.1 ms. It may be expected that much of significance will be learned as this type of preparation is studied with respect to the local processes and to comparison of nerves with component fibers of diverse proportions, excitability, accommodation and other properties.

Phenomena with probably the same basis have been observed in the central nervous system at the site of an interruption of the dorsal columns (12, 127) of the spinal cord. A very short delay is present in these cases also.

Artificial synapses in general have still much to offer, especially for quantitative study of those factors in transmission which are peculiarly under the control of the investigator in such preparations. They will continue to reward study even though natural synapses permitting observation of the unit junctions are rapidly becoming available.

Preparations for study of unit junctions. One of the most interesting results of recent work on various invertebrates has been the more or less incidental discovery of preparations permitting, with varying degrees of ease, the recording of activity of single synapses. The following are now available for this purpose. 1. The junction of primary sensory neurons from tactile receptors and second order ascending neurons in the last abdominal ganglion of the crayfish. 2. The similar synapse between cercal nerve afferents and giant central fibers in the last abdominal ganglion of the cockroach. 3. The synapses between medial as well as lateral giant fibers and segmental giant motor fibers in each abdominal ganglion of the crayfish. 4. That between the second order and third order giant fibers in the stellate ganglion of the squid. 5. Single nerve-muscle junctions isolated by dissection. It will be worthwhile to note the possibilities and limitations of these preparations.

1. Prosser's preparation has been mentioned above. While well suited to measurements of overall transmission properties and to experiments on altered external conditions, this preparation will probably not be convenient for recording the local signs of activity at the junction. Simultaneous activation of several

different ganglionic elements by each incoming impulse usually occurs, even though but one of these may be detected in the cord beyond the ganglion. Spontaneous firing of cells in the ganglion is continually going on under good conditions.

2. Pumphrey and Rawdon-Smith's preparation in the cockroach has many of the same advantages and disadvantages. The post-fibers being giant it may prove possible however to record local potentials relatively uncomplicated by small fiber activity, although the possibility of local reflex response (not appearing in the connective anteriorly) has not apparently been excluded. The existence of through fibers which are activated along with those that stop in the ganglion may prove a complication (but see 133). The principal limitations, however, are the difficulty of stimulating single pre-ganglionic units and the difficulty of confining the post-synaptic response to the same single unit for the duration of an experiment. There are a number of giant fibers, recruitment is well developed and the response at supraliminal stimulation is irregular.

3. Of quite a different character are the synapses in the crayfish between penultimate and final motor neurons of the giant system. These have been shown by Wiersma (154) to offer unique advantages. One asset, not yet fully exploited, is that the anatomy of the synapse, relative to most non-giant synapses, is quite accessible. It has been described for the related form, *Leander* (Holmes 72; see also Johnson, 85, on *Cambarus*). The four pre-fibers can be activated separately or in any combination, and in the root of the third peripheral nerve from each ganglion activity of a single post-fiber may be recorded. Some of the summation properties, the delay (0.7–1.0 ms.) and the fact of irreciprocal transmission have been reported but the possibilities for local response recording have not been explored. The circumstances are not ideal for single unit analysis owing to the presence of other junctions between the same pre-fibers and motor fibers of the first and second nerves.

4. The giant synapse in the stellate ganglion of the squid comes closer to the ideal for unit junction analysis. The intimate anatomy has been elegantly worked out by Young (162). A preliminary report of the physiologic properties has been published (Bullock, 37). There are two giant pre-fibers, whose synapses upon the post-fiber differ, and it is easy to stimulate one of these alone, although the other has not yet been so stimulated. A single post-fiber occurs in each of the stellar nerves and both pre- and post-fibers are of large diameter right up to the synapse. The fibers are in contact for nearly a millimeter. Leading directly from the ganglion, the local synaptic potential as well as both pre- and post-spikes are readily recorded.

Summation is not necessary; every presynaptic impulse is transmitted up to at least 300 per second. The delay is 1.2 ms. in a fresh preparation but can be doubled by fatigue. Transmission is polarized. Fatigue is readily brought on and maintained by stimulation at moderate frequencies, permitting study of the local response alone, while recovery and return of the spike can be induced at any time by reducing the frequency. Untransmitted impulses are thus seen to continue to exert an effect, preserving fatigue of the junction, apparently by

acting on the postsynaptic unit. Under such fatigue facilitation can be demonstrated. Refractory period of the local response may be absolute for a millisecond after which a response occurs but with an amplitude proportional to the interval, up to 2 or 3 ms. or more.[1] The synaptic potential when not intentionally depressed is estimated to rise at least three times higher than the threshold value for initiation of the propagated impulse. Its latency is not necessarily increased with fatigue but the height is reduced. The general picture forms an interesting comparison and contrast with the well-known work of Eccles (see 48 for references) on the synaptic potential recorded from many semi-synchronized synapses in the mammalian sympathetic ganglion. It is also very reminiscent of the artificial synapse of Arvanitaki (8) and the single neuro-myal junction of Kuffler (88).

5. The well-known work of Pantin (104 et ante) and Wiersma and van Harreveld (153 et ante) on the remarkable nerve-muscle relations in crustacea is being extended by study of the local events at the junction (Kuffler and Katz, 89, 86). Although it has apparently not been possible to isolate a single junction as was done in frog muscle, these authors are able to show the basis for the great dependence on facilitation. A single impulse arriving over an excitatory nerve fiber elicits only a small, non-propagated response comparable to the endplate potential described earlier in vertebrates. This local response grows with successive impulses at suitable frequencies (facilitation) and mechanical contraction accompanies it. Thus non-propagated, graded mechanical contraction is clearly demonstrated, in essentially intact limbs, as a normal mechanism, just as Wiersma (153) had predicted. Above a certain threshold of frequency propagated responses are said to occur. Confirming Wiersma and van Harreveld it is shown that impulses arriving over inhibitory nerve fibers can abolish both contraction and local potential produced by excitatory impulses. How this occurs can not yet be surmised but that it occurs subsequent to the action of the transmitter on the post-junctional membrane is indicated by the fact that facilitation continues during a period of inhibition, the response when inhibitor activity ceases immediately rising to the degree expected if no such period had supervened. The results do not offer any obvious aid in choosing between electrical and chemical transmitters, being compatible with either assumption. Further observations of this kind, particularly extensions of the technic to other muscles and other invertebrate groups will certainly repay effort.

Physiological neuronography. Dusser de Barenne (98) applied this term to the technic of mapping neuronal connections by following strychnine-induced spikes through the brain. It is equally appropriate for any technic designed to reveal intimate anatomy or functional organization by tracing electrical signs of nervous activity. This application of the methods of electrophysiology represents one of the most fruitful directions of investigation available to the comparative neurologist. Although well developed in vertebrate neurology, it has only begun to be exploited among invertebrates.

[1] In fresher preparations refractoriness of the local response may be absent at all intervals down to 0.2 ms., the shortest tested. Summation becomes prominent in these cases.

Prosser's (111) analysis of the pathways open to afferent volleys initiated in various parts of the crayfish is an excellent example of this approach, showing as it does many facts about functional neuroanatomy which would be impossible or extremely difficult to demonstrate by histological methods. Thus it was possible to specify that certain kinds of impulses, traveling in a given direction and resulting from a given type of stimulation, cross in the cord and others do not. The proper evaluation of the relations found largely awaits similar studies on related forms. The same author (113) mapped the area innervated by peripheral nerves in the earthworm, characterized the impulses elicited by various forms of stimulation and obtained evidence against the existence of a nerve net in the skin. Pringle (108) analyzed in a very successful way the reflex mechanism of the insect leg.

But virtually all other effort in this direction has concerned itself with the giant fiber systems found in widespread groups of both invertebrates and vertebrates. This is largely due to their technical accessibility, especially to electrical recording apparatus. The action potential in these fibers is so large that impulses in single fibers are easily recorded without elaborate dissection, indeed, in many cases from the external surface of the intact animal (139, 32, 35, 38). By tracing these impulses it is possible to reveal anatomical information such as the number of independent units, their extent (whether through-conduction or only short pathways), their anastomoses, contribution to peripheral nerves, connections with sensory and motor neurons and junctional properties.

Annelida. The first giant system from which action potentials were recorded was that in the earthworm. Eccles, Granit and Young (49), in a short note, reported that impulses were conducted in both directions and at a rather high speed. Earlier workers had concluded that the median fiber conducted posteriorly and the laterals anteriorly. Rushton (136, 138) and Bullock (31, 32) confirmed and extended these newer results. It was pointed out that the segmental septa which Stough (143) had shown to divide the fibers completely almost ten times every centimeter must not delay the impulse significantly. The problem of interpreting the septa is discussed above under "Junctional transmission." The earlier observations suggesting polarized conduction were explained by finding a topographic limitation of sensory connections to the giant fibers. The anatomically reported anastomosis between lateral fibers could be confirmed electrically. Motor connections have not been worked out but it is indicated that the three central giant fibers are probably not themselves efferent, that they elicit different muscular responses comprising the startle or withdrawal reflexes to stimuli applied to different parts of the body, and that the boundary between these regions must shift under different physiologic conditions (137).

Amassian and Floyd (5) describe repetitive firing of these fibers under certain conditions of galvanic stimulation.

The general features of the giant system are quite uniform among various oligochaetes but old anatomical evidence suggests that in the polychaetes a diversity of plan is to be found. This has begun to be exploited in recent studies (Bullock, 35, 38; Nicol and Young, 101). Many families of this large and diversi-

fied group lack giant fibers while others possess apparently fixed patterns consisting of from one unit to several dozen units. These may be compound or the process of a single cell, entirely central or efferent, long through-conduction paths or short, independent units, connected to sensory neurons throughout the body or only in restricted regions. All the through-conduction central fibers conduct both ways and agree in mediating a withdrawal response to startle stimuli. Giant fibers occur from less than 10 to more than 750 micra in diameter; apparent overall conduction rates, assuming no delays, fall in the range 1–10 m/s.

Arthropoda. Giant fiber systems in crustaceans, anatomically known from the accounts of Johnson (85), Lowe (92), and Holmes (72), have also been investigated physiologically only recently. Wiersma (152) reported unpolarized conduction and a certain pattern of interrelations between medial and lateral fibers as regards summation of effect on abdominal musculature in crayfish. Holmes, Pumphrey and Young (73) used these fibers in prawns to study conduction velocity as a function of fiber structure. Prosser and Altschule (121) report confirmation of Wiersma's general findings. The first extended account of a functional analysis of the crustacean giant system appeared in Wiersma's recent (154) report. He shows that the four distinct giant fibers all act on the same effectors, presumably through the same final motor neurons. No evidence is yet available of the difference in function between lateral and medial giants. Sensory connections are not known. As in the earthworm, the lateral fibers are connected together, but in the crayfish the connection is apparently synaptic, not anastomotic as has been assumed for the earthworm. Several interesting types of junctions between units of the system are discussed under "Junctional transmission."

The only studies on a comparable system in other arthropods are those of Pumphrey and Rawdon-Smith (125) and Roeder et al. (133) on the second order neurons in the auditory pathway from cercal receptors of the cockroach. These fibers ascend from the last abdominal ganglion, as through-conduction paths. Because of their high velocity and large diameter these authors use the term giant fibers but it may be questioned whether this is appropriate. In all other cases where the term is used a conspicuous discontinuity in size between the fibers in question and the next largest is recognizable. It is thus not used for the largest fibers in mammals though these are larger than some definite giant fibers in polychaetes, enteropneusts and others. Available evidence suggests that the large fibers in the cockroach as in other insects are the extremes of a continuous size spectrum. Nevertheless they permit recording and tracing of impulses and offer favorable opportunities for analysis of insect central nervous organization. The experiments of Pumphrey and Rawdon-Smith deal chiefly with the synapses in the last abdominal ganglion (see "Junctional transmission").

Mollusca. Until recently the only giant nerve fibers recognized in molluscs were those of decapods, although many early anatomists called attention to enormous nerve cells in various gastropods. Arvanitaki and Cardot have reported that the common snail *Helix* and the marine slug *Aplysia* display giant fibers in the interganglionic connectives (9, 41). Although an adequate descrip-

tive account of the anatomy is not given, a technic of extruding the cell bodies through their tube-like giant processes is presented. The chief interest of these authors has been the possibility of recording cell body potentials in isolated neuron somata. The potentials recorded in the connectives are remarkably long lasting (40 ms.) and slowly conducted (1 m/s). On present evidence it is difficult to orient the situation in these forms in terms of more familiar groups. Further elucidation of the gastropod plan of giant neurons will be awaited with interest.

The plan in decapods is at least as well known as that in any group, as a result of the anatomical work of Young (162). But, although the third order giant fiber from the squid has been the most studied giant preparation, as yet virtually no physiologic study has been made of the central organization of the system. The experiments of Young (161) on conduction through the stellate ganglion in the absence of cell bodies and of Bullock (37) on the properties of the synapse between second and third order giant fibers, in the same ganglion, are the only exceptions. The peripheral organization has been studied by Prosser and Young (123) and Pumphrey and Young (126) who have shown the neuromuscular relations and the significance of conduction velocity.

On the whole, the investigations made thus far upon giant fiber systems have been almost entirely exploratory. The preparations thus made available offer great opportunities for future work both upon functional organization at the level of comparative physiology and upon fundamental properties of units at the general physiologic level. Still other preparations are possible as well. Giant systems are known in many groups, even as low as the enteropneusts (30, 33) whose nervous system is very primitive. The high speed conduction tract in the mesenteries of anthozoans (*Calliactis*, Pantin, 103), overlapping as they do the lowest conduction rates observed in known giant fibers of polychaetes (38) suggest that possibly something very like a giant system is present already in the lowest phylum having a differentiated nervous system. Forms such as the stomatopod, *Squilla*, would seem to offer favorable circumstances for special studies, owing to large size and transparency of exoskeleton over the nerve cord; they have been found to possess good giant spikes (Bullock, unpub.). The need of further work on the giants of lower vertebrates (Mauthner and Müller fibers) may also be mentioned. These classical objects of anatomical study, source of some of our best cytologic knowledge of the vertebrate synapse, have only received their first electrophysiologic study (58). Of all giant systems least is known concerning the function in these vertebrate ones, much studied as they have been by experimental embryologic methods. It would seem that much of interest could be learned by recording from the fibers in life under conditions of physiologic, as well as direct, stimulation, but the necessity of finding particularly favorable species must be emphasized.

Spontaneous Central Activity. As is well known from the work of Adrian (1, 2), Rijlant (128), Prosser (109, 116), Bonnet and Bremer (19), Bonnet (17, 18), Roeder (134, 131), activity occurs in the central nervous system of arthropods in the absence of apparent stimulation. This has been extended to annelids (62, 63, 99, 34) and molluscs (26, 34) but as yet to no other groups besides verte-

brates. This last does not mean other groups have been found to lack such activity. Spontaneous activity has been found in every animal suitably tested. Like that in vertebrates, it is not fundamentally altered by deafferentation and persists in isolated ganglia or pieces thereof for long periods under conditions which need not be regarded as abnormal or excitatory. The consistent finding is striking that the general pattern of activity is much the same in the intact and probably excited animal, the intact and probably unexcited animal and the isolated ganglia after frank injury discharge ceases.

Interest in this type of nervous function has centered around certain problems: 1, spontaneity as a general phenomenon and its relation to rhythmic pacemakers; 2, the influence of environmental agents; 3, the bearing of such activity in invertebrates upon the interpretation of brain waves as seen in man, and 4, the significance of special types of pattern such as "visual rhythms."

Studies on the first of these have been thus far largely descriptive and have demonstrated the widespread occurrence of the phenomenon. It is present in the quiescent moth pupa (34) during a time when many of the neurons are dedifferentiating, others differentiating in a general reorganization of connections, sensory and motor functions are at a low ebb, appendages are absent and the musculature is largely dedifferentiated. It is apparent in small fragments of isolated ganglia, as in the cardiac ganglion of *Limulus* (6, 39). Even some peripheral sensory neurons exhibit spontaneous discharge (59, 94, 140, 53). But it is absent in many neurons under normal conditions; for example, giant neurons of earthworm, crayfish and squid. Experimental analysis has begun, with the study of agents which can alter the pace or initiate it in inactive cells and fibers (electrical polarization, 11, 50, 7, 39; oxidative control, 122; ions, 117, 118, 120, 131; temperature, 112; drugs, 20, 134; general, 50, 18, 24). Details cannot be given here but it is apparent that many factors contribute to the determination of frequency of a pacemaker and that under appropriate conditions any neuron can be made to discharge rhythmically in the presence of a steady state environment, in the manner of a relaxation oscillator. It is therefore a question of the individual properties of the neuron, its thresholds and time constants, whether and how active it will be under particular "normal" conditions. The phenomenon of spontaneity is, of course, not thereby explained but the quantitative definition of the difference between spontaneously active and inactive neurons is begun.

The relation of invertebrate manifestations of spontaneity to the interpretation of vertebrate brain waves has been discussed elsewhere (34). It was pointed out that comparative studies reveal a high degree of uniformity in the general pattern, especially the general position of the frequency spectrum of brain waves in all vertebrates. The uniformity is such that the smooth, slow (chiefly 1–30 per sec.) character of the activity obtains alike (though not without some significant differences) for intact human brain, deafferented cortex, isolated basal ganglia, the non-corticated frog brain and isolated fragments of the frog olfactory lobe as small as 0.1 mgm. "This, together with other evidence, suggests that brain waves should be looked upon not as a sign of the higher aspects of nervous activ-

ity but as a reflection of some basic, primitive common denominator of the brain of frogs and men." It is this general agreement in character among vertebrates that gives significance to the comparison with invertebrates, for a similar agreement exists among them, but the common pattern of invertebrates is strikingly different from that of vertebrates. The former are characterized by fast, spikey activity, many spikes having the time relations of single axon action currents and the dominant frequency being probably twenty times that of vertebrates. Discussion of the evidence leads to the conclusions that the prominence of spikes in one group and their virtual absence in the other is not readily attributed to size, or to complexity or synchronization, that slow waves comparable to vertebrate brain waves are present in many invertebrate ganglia (17, 18) and that both here and in vertebrate brain waves they are probably not explained as a resultant of impulse action currents but rather as a separate form of neuronal activity—a slowly oscillating change of state or beating (Gerard, 55) of masses of nerve cells in unison. In support of the last are such facts as these: the slow and spikey components in an insect brain can be altered by environmental agents separately, each can exist in the virtual absence of the other, in rhythmically discharging ganglia like the cardiac ganglion of *Limulus*, each discharge not only includes both components but the slow waves may initiate the discharge, preceding the first spike. Further evidence, including slow wave propagation in the absence of synaptic pathways, has been presented by Gerard (55) from studies on vertebrates and discussed in Bullock (34). The far-reaching implications of the assumption that neurons can affect each other by means distinct from classical impulses in synaptic pathways are obvious (see Bremer, 24). It seems apparent that the comparative approach to brain waves and spontaneous activity generally has significant contributions to make to the basic understanding of these aspects of neural function.

An important special case of slow activity has been extensively studied (Adrian, 3, 4; Crescitelli, 45; Jahn, Wulff, 80; Bernhard, 13) in the insect. Here rhythmic potentials occur which are not strictly spontaneous in that they are related to visual stimulation, even though they often occur for considerable periods after stimulation has ceased. Much work has gone into the descriptive analysis of the visual rhythms and the localization of the activity but as yet little work has been directed toward the problem of their basic nature. If their strong resemblance to vertebrate brain waves is more than superficial, as seems likely, they should represent useful material for experiments on the meaning of brain waves. Adrian (4) evidently regarded the oscillatory potential of the insect visual system as a composite of impulse discharges, but evidence has been cited (34 and below, p. 657) that although spikes may be synchronized with a certain phase of the slow waves the latter cannot be explained as an envelope or resultant of nothing but impulse spikes.

Sense Organs. As has been abundantly demonstrated in vertebrate physiology, one of the most successful applications of electrophysiologic methods is in the identification and characterization of sensory structures. Examples of this approach on invertebrate material serve to emphasize its possibilities. Although

some of these cases have been reviewed elsewhere, it seems desirable to bring them together with more recent ones to underline the opportunities for further work.

Prosser (110) discovered by action current recording the presence of a primitive photoreceptor system in the last abdominal ganglion of the crayfish. The properties of such cells are of great interest as they are apparently still undifferentiated anatomically from ordinary nerve cells (Welsh, 148). There are evidently only a few units but they can play a significant rôle in the activity of the animal deprived of eyes (148). Latency is long, maximum response is attained only after some seconds and may even occur after the light is turned off, adaptation occurs but slightly and slowly, activity continues for some seconds after the stimulus is removed. This preparation is well worth further attention.

Other primitive light receptors have been studied in the pelecypods. Hartline (68) found that one group of cells in the eye of the scallop responds, much like a large group of the elements of the vertebrate eye, with a burst of impulses at the "on," reaching highest frequency almost immediately (in contrast to the caudal photoreceptor of crayfish) and adapting greatly, ceasing promptly at "off"—with rarely an after discharge. Another group of cells also believed to be primary sense cells are "off"-responding like another large group of elements in the vertebrate eye—silent during illumination and firing at its cessation or reduction, to a degree dependent on the degree of illumination, suppressed promptly by reillumination, and often synchronized into rhythmic bursts. An especially notable feature of these two types of elements in the scallop eye is that they are anatomically separate and send their axons through separate nerves. Piéron and Segal (106) recorded slow waves and nerve action currents attributable to the scattered photoreceptors in the skin of the siphons of the clam *Mya*. On and off responding systems had distinct properties.

More advanced light receptors, forming well developed eyes, have been studied surprisingly little considering the elegance of the technics and the accessibility of the structures. Analysis of properties of units especially has lagged. The classical work of Hartline and his collaborators (57, 70, 69 et ante) on the single units in *Limulus* still stands unique. A point of special interest in one of the most recent contributions of this group (157) has not received the attention it deserves. Although off receptors are said to be absent from the *Limulus* eye, an off response is found in the ganglion where the primary sense cells synapse with second order neurons. The off response is as readily elicited by artificial stimulation of the nerve carrying the primary afferents as by photic stimulation of the eye: upon cessation of the afferent volley to the ganglion certain neurons therein discharge. This phenomenon has been heretofore associated with receptor elements only; the possibility that it may be a normal mechanism of central synaptic activation is not an accepted part of current neurophysiologic teaching. Studies directed toward the questions (a) is this in fact a generally used central excitatory mechanism? and (b) how can it be understood on any of the existing theories of synaptic transmission? are urgently needed.

A considerable body of work has accumulated dealing with the large slow po-

tentials associated with the illumination of photoreceptor elements in insects and cephalopods. These "retinal" potentials resemble in general the responses in vertebrate eyes and in some forms direct identification of waves with those of vertebrates appears possible (146). In other cases agreement has not been reached on the component waves even in closely related animals (13, 145). Analysis of the waves has led to the assignment of some to the optic ganglion and others to the receptors alone (80, 13). It has shown how the pattern varies with intensity and the state of adaptation (75, 43, 159, 67), with a remarkable diurnal rhythm in sensitivity (77, 79, 82, 159) and with temperature (81). In spite of the considerable effort, we are still much in need of comprehensive analytical work towards the definition of common denominators. And thus far the electrical events have not been successfully identified with any chemical or excitatory processes. Significant relations with nerve impulses are indicated, however, in the facts (1) that certain of the slow waves always begin before nerve fiber spikes appear, as is true in vertebrates (Adrian, 2; Bullock, unpublished observations), (2) at a certain phase of the slow "off" deflection all spike activity, even the spontaneous fraction, may be inhibited (Bernhard, 13) and (3) spike activity during illumination tends to be grouped into bursts in phase with the rhythmic oscillatory slow waves (4, 13, 36).

The distinction between two groups of fibers in the peripheral nerves of the earthworm, whose end organs are sensitive to two different kinds of stimulation (113), represents one of the most primitive cases of differentiated sensory structures physiologically recognized. It may be hoped that the method will be applied at even lower phyelogenetic levels. These findings doubtless represent the lowest known proprioceptors. Their properties and the possible occurrence of stretch receptors in other forms, including nemerteans and polychaetes, should be examined with special reference to the unique degree of extensibility, both active and passive, in these animals.

Proprioceptors have been identified elsewhere among invertebrates only in insects. As a result of the work of Pringle (107) special receptors in the cockroach can be designated as position sense organs though they are also receptive of other mechanical stimuli. They are of special interest because of their mechanism and properties and because of a peculiar importance in DDT poisoning. Pringle has recognized two kinds, campaniform sensilla and hair sensilla. Each is distributed in specific patterns on the appendages and body, strategically located to detect stresses in the exoskeleton and movement of folds in intersegmental membranes, respectively.

The campaniform sensilla appear to be specifically sensitive to compressive forces in the cuticle and each end organ, indeed each group of sensilla is oriented so as to detect only components along a certain axis. Their properties are similar to some types of vertebrate proprioceptors but a difference of probably broad significance is that they respond not to tension in individual muscles but to the integrated result of several muscles, the force of gravity and the position of the member in relation to the force. Here is yet another case of the peripheral integration seen in so many invertebrate sensory and motor relations. A point

of anatomical interest is the strong probability that the nerve fibers from several end organs join to form a single afferent fiber in the peripheral nerve—as is likely also in annelids where many more primary sense cells than fibers in the segmental nerves are found (113). It should be emphasized that there are probably many other receptors acting as proprioceptors, for the anatomy of insects is such that most mechano-receptors (tactile hairs, chordotonal organs, etc.) must respond to touch, pressure, muscle tension, gravity and movement, rarely being able to distinguish one form of stimulus from others (see Wigglesworth, 155).

Roeder (132) and Roeder and Weiant (135) have shown these structures to be especially susceptible to DDT and probably the first organs to respond to this poison, although no definite relation between its action on these receptors and its lethality can be stated. In great dilution DDT evokes prolonged trains of impulses in the nerve fibers from the proprioceptors, bombarding the central nervous system with an intensity that might well explain some of the observed motor effects (147).

Sound receptors have been studied electrically in insects by Wever et al. (151, 150) and Pumphrey and Rawdon-Smith (124), whose work is reviewed by Wigglesworth (155).

Among the more conspicuous opportunities in the area of sensory physiology the following may be mentioned: confirmation and characterization of the gyroscopic organs of equilibrium in the halteres of dipterans and others (Wigglesworth, 156); demonstration of the distinctness, or lack of it, of receptors for special forms of stimuli—thus, touch and pressure, pressure and muscle tension, heat and cold, nociceptive and heat or cold, etc. (Frings, 52), subdivisions of the chemical sense (51, 155); single unit analysis of photoreceptors for the direct demonstration of color reception in insects and others—in the manner of Granit's already classical work on vertebrates (see 60).

If any general conclusions are to be drawn from the group of studies here cited they would certainly emphasize anew the pregnancy of the comparative approach both for providing leads, clues and perspectives of value in strictly mammalian neurophysiology and for providing material uniquely suited to many fundamental experiments in general neurology. Thus far work in this area has gone little beyond exploration; the principal opportunities are still ahead.

REFERENCES

(1) Adrian, E. D. The activity of the nervous system of the caterpillar. J. Physiol. 70: 34, 1930.

(2) Adrian, E. D. Potential changes in the isolated nervous system of Dytiscus marginalis. J. Physiol. 72: 132, 1931.

(3) Adrian, E. D. The activity of the optic ganglion of Dytiscus marginalis. J. Physiol. 75: 26P, 1932.

(4) Adrian, E. D. Synchronized reactions in the optic ganglion of Dytiscus. J. Physiol. 91: 66, 1937.

(5) Amassian, V. E. and W. F. Floyd. Repetitive discharge of giant nerve fibers of the earthworm. Nature 157: 412, 1946.

(6) Armstrong, F., M. Maxfield, C. L. Prosser and G. Schoepfle. Analysis of the electrical discharge from the cardiac ganglion of Limulus. Biol. Bull. 77: 327, 1939.

(7) Arvanitaki, A. Propriétés rythmiques de la matière vivante. Variations graduées de la polarisation et rythmicités. Paris. Hermann, 1938.

(8) Arvanitaki, A. Effects evoked in an axon by the activity of a contiguous one. J. Neurophysiol. 5: 89, 1942.

(9) Arvanitaki, A. and H. Cardot. Observations sur la constitution des ganglions et conducteurs nerveux et sur l'isolement du soma neuronique vivant chez les mollusques gastropodes. Bull. Hist. Appl. Physiol. et Path. 18: 134, 1941.

(10) Ashworth, J. H. The giant nerve cells and fibres of Halla parthenopeia. Philos. Trans. Royal Soc., London B 200: 427, 1909.

(11) Auger, C. and A. Fessard. Observations complementaries sur un phénomène de contractions rythmées provoquées par excitation galvanique chez certains insects. C. R. Soc. Biol. 101: 897, 1929.

(12) Barron, D. H. Central course of "recurrent sensory discharges." J. Neurophysiol. 3: 403, 1940.

(13) Bernhard, C. G. Isolation of retinal and optic ganglion response in the eye of Dytiscus. J. Neurophysiol. 5: 32, 1942.

(14) Bishop, G. H. The relation of bioelectric potentials to cell functioning. Ann. Rev. Physiol. 3: 1, 1941.

(15) Bishop, G. H. The structural identity of the pain spot in human skin. J. Neurophysiol. 7: 185, 1944.

(16) Bishop, G. H. Nerve and synaptic conduction. Ann. Rev. Physiol. 8: 355, 1946.

(17) Bonnet, V. Contribution a l'étude du système nerveux ganglionnaire des crustacés. Arch. Internat. Physiol. 47: 397, 1938.

(18) Bonnet, V. L'activite rythmique de la cellule nerveuse et ses modifications. Lyons, Soc. Anon. de l'Imprimerie A. Roy, 1941.

(19) Bonnet, V. and F. Bremer. Étude oscillographique de l'activité électrique spontaneé de la cellule nerveuse des crustacés. C. R. Soc. Biol. 127: 798, 1938.

(20) Bonnet, V. and F. Bremer. Action de la strychnine et de l'acetylcholine sur la rythmicité neuronique chez les crustacés. C. R. Soc. Biol. 127: 804, 1938.

(21) Botsford, E. F. The effect of physostigmine on the responses of earthworm body wall preparations to successive stimuli. Biol. Bull. 80: 299, 1941.

(22) Boule, L. Recherches sur le système nerveux central normal du Lombric. Le Nevraxe 10: 16, 1908.

(23) Bozler, E. Untersuchungen über das Nervensystem der Coelenteraten. I. Teil: Kontinuitat oder Kontakt. Ztschr. Zellforsch. Micr. Anat. 5: 244, 1927.

(24) Bremer, F. L'activité "spontanée" des centres nerveus. Bull. Acad. Roy. Med. Belgique, ser. 6 9: 148, 1944.

(25) Brink, F., D. W. Bronk and M. G. Larrabee. Chemical excitation of nerve. Ann. N. Y. Acad. Sci. 47: 457, 1946.

(26) Bronk, D. W., J. Z. Young and R. W. Gerard. Electrical activity of isolated stellar ganglion of Loligo. Unpublished observations cited in Gerard, Cold Spring Harbor Symp. 4: 293, 1936.

(27) Bullock, T. H. The functional organization of the nervous system of Enteropneusta. Biol. Bull. 79: 91, 1940.

(28) Bullock, T. H. The existence of unpolarised synapses. Anat. Rec. 78: 67, 1940.

(29) Bullock, T. H. Neuromuscular facilitation in scyphomedusae. J. Cell. Comp. Physiol. 22: 251, 1943.

(30) Bullock, T. H. The giant nerve fiber system in balanoglossids. J. Comp. Neurol. 80: 355, 1944.

(31) Bullock, T. H. Oscillographic studies on the giant nerve fiber system in Lumbricus. Biol. Bull. 87: 159, 1944.

(32) Bullock, T. H. Functional organization of the giant fiber system of Lumbricus. J. Neurophysiol. 8: 55, 1945.

(33) Bullock, T. H. Anatomical organization of the nervous system of Enteropneusta. Quart. J. Micr. Sci. 86: 55, 1945.

(34) Bullock, T. H Problems in the comparative study of brain waves. Yale J Biol Med. 17: 657, 1945.
(35) Bullock, T. H. Organisation of the giant nerve fiber system in Neanthes virens. Biol. Bull. 89: 185, 1945.
(36) Bullock, T. H. The double beam cathode ray tube in biological research. Electronics 19: 103, 1946.
(37) Bullock, T. H. A preparation for the physiological study of the unit synapse. Nature 158: 555, 1946.
(38) Bullock, T. H. Physiological mapping of giant nerve fiber systems in polychaete annelids. Physiol. Comp. et ecol., in press.
(39) Bullock, T. H., H. S. Burr and L. F. Nims. Electrical polarization of pacemaker neurons. J. Neurophysiol. 6: 85, 1943.
(40) Bullock, T. H., H. Grundfest, D. Nachmansohn and M. Rothenberg. Effect of di-isopropyl fluorophosphate (DFP) on action potential and cholinesterase of nerve. II. J. Neurophysiol. 10: 63, 1947.
(41) Cardot, H. and A. Arvanitaki. Données sur les caractéristiques de l'activité électrique du soma neuronique. Schweiz. Med. Wchnschr. 71: 395, 1941.
(42) Carter, G. S. A general zoology of the invertebrates. New York, Macmillan, 1940.
(43) Crescitelli, F. and T. L. Jahn. The electrical response of the dark adapted grasshopper eye to various intensities of illumination and to different qualities of light. J. Cell. Comp. Physiol. 13: 105, 1939.
(44) Crescitelli, F. and T. L. Jahn. The effect of temperature on the electrical response of the grasshopper eye. J. Cell. Comp. Physiol. 14: 13, 1939.
(45) Crescitelli, F. and T. L. Jahn. Oscillatory electrical activity from the insect compound eye. J. Cell. Comp. Physiol. 19: 47, 1942.
(46) Curtis, H. J. and K. S. Cole. Nerve: excitation and propagation. In Medical physics, O. Glasser, ed., Chicago, Yearbook Pub., 1944.
(47) Eccles, J. C. An electrical hypothesis of synaptic and neuro-muscular transmission. Nature 156: 680, 1945.
(48) Eccles, J. C. An electrical hypothesis of synaptic and neuro-muscular transmission. Ann. N. Y. Acad. Sci. 47: 429, 1946.
(49) Eccles, J. C., R. Granit and J. Z. Young. Impulses in the giant fibres of earthworms. J. Physiol. 77: 23P, 1933.
(50) Fessard, A. Propriétés rythmiques de la matière vivante. I. Nerfs isolés nerfs myélinisés. II. Nerfs isolés—nerfs non-myélinisés. Paris, Hermann, 1936.
(51) Frings, H Gustatory rejection thresholds for the larvae of the cecropia moth, Samia cecropia (Linn.). Biol. Bull. 88: 37, 1945.
(52) Frings, H. The reception of mechanical and thermal stimuli by caterpillars. J. Exp. Zool. 99: 115, 1945.
(53) Galambos, R. and H. Davis. Inhibition of activity in single auditory nerve fibers by acoustic stimulation. J. Neurophysiol. 7: 287, 1944.
(54) Gamble, F. W. and J. H. Ashworth. The anatomy and classification of the Arenicolidae, with some observations on the post-larval stages. Quart. J. Micr. Sci. 43: 419, 1900.
(55) Gerard, R. W. The interaction of neurones. Ohio J. Sci. 41: 160, 1941.
(56) Gerard, R. W. Nerve metabolism and function. Ann. N. Y. Acad. Sci. 47: 575, 1946.
(57) Graham, C. H. and H. K. Hartline. The response of single visual sense cells to lights of different wave lengths. J. Gen. Physiol. 18: 917, 1935.
(58) Graham, H. T. and J. L. O'Leary. Fast central fibers in fish. Properties of Mauthner and Müller fibers of medullospinal system. J. Neurophysiol. 4: 224, 1941.
(59) Granit, R. Rotation of activity and spontaneous rhythms in the retina. Acta Physiol. Scand. 1: 370, 1941.
(60) Granit, R. The electrophysiological analysis of the fundamental problem of colour reception. Proc. Physical Soc. 57: 447, 1945.

(61) GRANIT, R. AND C. R. SKOGLUND. Facilitation, inhibition and depression at the "artificial synapse" formed by the cut end of a mammalian nerve. J. Physiol. **103**: 435, 1945.
(62) GRAY, J. AND H. W. LISSMANN. Studies in animal locomotion. VII. Locomotory reflexes in the earthworm. J. Exper. Biol. **15**: 506, 1938.
(63) GRAY, J., H. W. LISSMANN AND R. J. PUMPHREY. The mechanism of locomotion in the leech (Hirudo medicinalis Ray). J. Exper. Biol. **15**: 408, 1938.
(64) GRUNDFEST, H. Bioelectric potentials in the nervous system and in muscle. Ann. Rev. Physiol. **9**: 477, 1947.
(65) HAMAKER, J. J. The nervous system of Nereis virens Sars. A study in comparative neurology. Bull. Mus. Harvard Coll. **32**: 89, 1898.
(66) HANSTRÖM, B. Vergleichende Anatomie des Nervensystems der Wirbellosen Tiere. Berlin, Springer, 1928.
(67) HARTLINE, H. K. The dark adaptation of the eye of Limulus, as manifested by its electric response to illumination. J. Gen. Physiol. **13**: 379, 1930.
(68) HARTLINE, H. K. The discharge of impulses in the optic nerve of Pecten in response to the illumination of the eye. J. Cell. Comp. Physiol. **11**: 465, 1938.
(69) HARTLINE, H. K. Sense organs. Ann. Rev. Physiol. **4**: 445, 1942.
(70) HARTLINE, H. K. AND R. McDONALD. Dark adaptation of single visual sense cells. Am. J. Physiol. **133**: 321P, 1941.
(71) HODGKIN, A. L. AND W. A. H. RUSHTON. The electrical constants of a crustacean nerve fibre. Proc. Roy. Soc. London B **133**: 444, 1946.
(72) HOLMES, W. The giant myelinated nerve fibres of the prawn. Phil. Trans. Roy Soc. London B **231**: 293, 1942.
(73) HOLMES, W., R. J. PUMPHREY AND J. Z. YOUNG. The structure and conduction velocity of the medullated nerve fibres of prawns. J. Exper. Biol. **18**: 50, 1941.
(74) HYMAN, L. H. The invertebrates: Protozoa through Ctenophora. New York, McGraw-Hill, 1940.
(75) JAHN, T. L. AND F. CRESCITELLI. The electrical response of the grasshopper eye under conditions of light and dark adaptation. J. Cell. Comp. Physiol. **12**: 39, 1938.
(76) JAHN, T. L. AND F. CRESCITELLI. The electrical response of the Cecropia moth eye. J. Cell. Comp. Physiol. **13**: 113, 1939.
(77) JAHN, T. L. AND F. CRESCITELLI. Diurnal changes in the electrical response of the compound eye. Biol. Bull. **78**: 42, 1940.
(78) JAHN, T. L. AND F. CRESCITELLI. Electrical oscillation from insect eyes. Am. J. Physiol. **133**: 339, 1941.
(79) JAHN, T. L. AND V. J. WULFF. Influence of a visual diurnal rhythm on flicker response contours of Dytiscus. Proc. Soc. Exper. Biol. Med. **48**: 660, 1941.
(80) JAHN, T. L. AND V. J. WULFF. Allocation of electrical responses from the compound eye of grasshoppers. J. Gen. Physiol. **26**: 75, 1942.
(81) JAHN, T. L. AND V. J. WULFF. Effect of temperature upon the retinal action potential. J. Cell. Comp. Physiol. **21**: 41, 1943.
(82) JAHN, T. L. AND V. J. WULFF. Electrical aspects of a diurnal rhythm in the eye of Dytiscus fasciventris. Physiol. Zool. **16**: 101, 1943.
(83) JAHN, T. L. AND V. J. WULFF. The spectral sensitivity of Dytiscus fasciventris. Anat. Rec. **96**: 11, 1946.
(84) JASPER, H. W. AND A. M. MONNIER. Transmission of excitation between excised non-myelinated nerves. An artificial synapse. J. Cell. Comp. Physiol. **11**: 259, 1938.
(85) JOHNSON, G. E. Studies on the functions of the giant fibers of crustaceans, with special reference to Cambarus and Palaemonetes. J. Comp. Neurol. **42**: 19, 1926.
(86) KATZ, B. AND S. W. KUFFLER. Excitation of the nerve-muscle system in crustacea. Proc. Roy. Soc. London B **133**: 374, 1946.

(87) KATZ, B. AND O. H. SCHMIDT. Electric interaction between two adjacent nerve fibers. J. Physiol. **97**: 471, 1940.
(88) KUFFLER, S. W. Further study on transmission in an isolated nerve-muscle fibre preparation. J. Neurophysiol. **5**: 309, 1942.
(89) KUFFLER, S. W. AND B. KATZ. Inhibition at the nerve-muscle junction in crustacea. J. Neurophysiol. **9**: 337, 1946.
(90) LEWIS, M. Studies on the central and peripheral nervous systems of two polychaete annelids. Proc. Am. Acad. Arts. Sci. **33**: 225, 1898.
(91) LIVINGSTON, W. K. Pain mechanisms. New York, Macmillan, 1943.
(92) LOWE, E. On the anatomy of a marine copepod, Calanus finmarchicus (Gunnerus). Trans. Roy. Soc. Edinburgh **58**: 561, 1936.
(93) LOWENSTEIN, O. A method of physiological assay of pyrethrum extracts. Nature **150**: 760, 1942.
(94) LOWENSTEIN, O. AND A. SAND. The activity of the horizontal semicircular canal of the dogfish, Scyllium caniculae. J. Exper. Biol. **13**: 416, 1936.
(95) MARRAZZI, A. S. AND R. LORENTE DE NO. Interaction of neighboring fibres in myelinated nerve. J. Neurophysiol. **7**: 83, 1944.
(96) McCONNELL, C. H. The development of the ectodermal nerve net in the buds of Hydra. Quart. J. Micr. Sci. **75**: 495, 1932.
(97) McCULLOUCH, W. S. Irreversibility of conduction in the reflex arc. Science **87**: 65, 1938.
(98) McCULLOUCH, W. S. Cortico-cortical connections, in "The Precentral Motor Cortex," P. C. Bucy, ed., Urbana, Univ. Ill. Press, 1944.
(99) MOORE, A. R. AND W. BRADWAY. The significance of action potentials in the isolated nerve cord of the earthworm. J. Cell. Comp. Physiol. **25**: 181, 1945.
(100) NACHMANSOHN, D. Chemical mechanism of nerve activity. Ann. N. Y. Acad. Sci. **47**: 395, 1946.
(101) NICOL, J. A. C. AND J. Z. YOUNG. Giant nerve fibres of Myxicola infundibulum (Grube). Nature **158**: 167, 1946.
(102) OGAWA, F. The nervous system of earthworm (Pheretima communissima) in different ages. Sci. Reports, Tohuku Imp. Univ., 4th ser. **13**: 395, 1939.
(103) PANTIN, C. F. A. The nerve net in the Actinozoa. IV. Facilitation and the "Staircase." J. Exper. Biol. **12**: 389, 1935.
(104) PANTIN, C. F. A. Junctional transmission of stimuli in the lower animals. Proc. Roy. Soc. B **123**: 397, 1937.
(105) PARKER, G. H. The elementary nervous system. New York, Macmillan, 1919.
(106) PIÉRON, H. AND H. SEGAL. Les manifestations électriques de l'excitation lumineuse chez la Mye. C. R. Soc. Biol. Paris **130**: 47, 1939.
(107) PRINGLE, J. W. S. Proprioception in insects. I. A new type of mechanical receptor from the palps of the cockroach. II. The action of the campaniform sensilla on the legs. III. The function of the hair sensilla at the joints. J. Exper. Biol. **15**: 101, 114, 467, 1938.
(108) PRINGLE, J. W. S. The reflex mechanism of the insect leg. J. Exper. Biol. **17**: 8, 1940.
(109) PROSSER, C. L. Action potentials in the nervous system of the crayfish. I. Spontaneous impulses. J. Cell. Comp. Physiol. **4**: 185, 1934.
(110) PROSSER, C. L. Action potentials in the nervous system of the crayfish. II. Responses to illumination of the eye and caudal ganglion. J. Cell. Comp. Physiol. **4**: 363, 1934.
(111) PROSSER, C. L. Action potentials in the nervous system of the crayfish. III. Central responses to proprioceptive and tactile stimulation. J. Comp. Neurol. **62**: 495, 1935.
(112) PROSSER, C. L. Action potentials in the nervous system of the crayfish. IV. Influence of temperature on nerve impulses arising "spontaneously" in abdominal ganglia. J. Gen. Physiol. **19**: 65, 1935.

(113) PROSSER, C. L. Impulses in the segmental nerves of the earthworm. J. Exper. Biol. 12: 95, 1935.
(114) PROSSER, C. L. A preparation for the study of single synaptic junctions. Am. J. Physiol. 113: 108, 1935.
(115) PROSSER, C. L. Action potentials in the nervous system of the crayfish. V. Temporal relations in presynaptic and postsynaptic responses. J. Cell. Comp. Physiol. 7: 95, 1935.
(116) PROSSER, C. L. Rhythmic activity in isolated nerve centers. Cold Spring Harbor Symp. Quant. Biol. 4: 339, 1936.
(117) PROSSER, C. L. Evidence for chemical control of "spontaneous" activity of isolated ganglia. Am. J. Physiol. 123: 165, 1938.
(118) PROSSER, C. L. Effects of salts upon spontaneous activity in the nervous system of crayfish. J. Cell. Comp. Physiol. 15: 55, 1940.
(119) PROSSER, C. L. Action potentials in the nervous system of the crayfish. Effects of drugs and salts upon synaptic transmission. J. Cell. Comp. Physiol. 16: 25, 1940.
(120) PROSSER, C. L. Effects of ions upon isolated nerve centers. Am. J. Physiol. 133: 417, 1941.
(121) PROSSER, C. L. The physiology of nervous systems of invertebrate animals. Physiol. Rev. 26: 337, 1946.
(122) PROSSER, C. L. AND C. C. BUEHL. Oxidative control of "spontaneous" activity in the nervous system of the crayfish. J. Cell. Comp. Physiol. 14: 287, 1939.
(123) PROSSER, C. L. AND J. Z. YOUNG. Responses of muscles of the squid to repetitive stimulation of the giant nerve fibers. Biol. Bull. 73: 237, 1937.
(124) PUMPHREY, R. J. AND A. F. RAWDON-SMITH. Hearing in insects: the nature of the response of certain receptors to auditory stimuli. Proc. Roy. Soc. London B 121: 18, 1936.
(125) PUMPHREY, R. J. AND A. F. RAWDON-SMITH. Synaptic transmission of nervous impulses through the last abdominal ganglion of the cockroach. Proc. Roy. Soc. London B 122: 106, 1937.
(126) PUMPHREY, R. J. AND J. Z. YOUNG. The rates of conduction of nerve fibres of various diameters in cephalopods. J. Exper. Biol. 15: 453, 1938.
(127) RENSHAW, B. AND P. O. THERMAN. Excitation of intraspinal mammalian axons by nerve impulses in adjacent axons. Am. J. Physiol. 133: 96, 1941.
(128) RIJLANT, P. L'activité du ganglion cardiaque de la Limule polyphème. C. R. Soc. Biol. 108: 828, 1931.
(129) ROEDER, K. D. Synchronized activity in the optic and protocerebral ganglia of the grasshopper, Melanoplus femur-rubrum. J. Cell. Comp. Physiol. 14: 299, 1939.
(130) ROEDER, K. D. The origin of visual rhythms in the grasshopper, Melanoplus femur-rubrum. J. Cell. Comp. Physiol. 16 (suppl): 1, 1940.
(131) ROEDER, K. D. The effect of potassium and calcium on the spontaneous activity of the isolated crayfish nerve cord. J. Cell. Comp. Physiol. 18: 1, 1941.
(132) ROEDER, K. D. The action of DDT on the campaniform organs of the cockroach. Anat. Rec. 96: 3, 1946.
(133) ROEDER, K. D., N. K. KENNEDY AND E. A. SAMSON. Synaptic conduction to giant fibers of the cockroach and the action of anticholinesterases. J. Neurophysiol. 10: 1, 1947.
(134) ROEDER, K. D. AND S. ROEDER. Electrical activity in the isolated ventral nerve cord of the cockroach. I. The action of pilocarpine, nicotine, eserine and acetylcholine. J. Cell. Comp. Physiol. 14: 1, 1939.
(135) ROEDER, K. D. AND E. H. WEIANT. The site of action of DDT in the cockroach. Science 103: 304, 1946.
(136) RUSHTON, W. A. H. Action potentials from the isolated nerve cord of the earthworm. Proc. Roy. Soc. London B 132: 423, 1945.

(137) RUSHTON, W. A. H. Motor responses from giant fibres in the earthworm. Nature 156: 109, 1945.

(138) RUSHTON, W. A. H. Reflex conduction in the giant fibres of the earthworm. Proc. Roy. Soc. B 133: 109, 1946.

(139) RUSHTON, W. A. H. AND H. B. BARLOW. Single-fibre response from an intact animal. Nature 152: 597, 1943.

(140) SAND, A. The function of the ampullae of Lorenzini with some observations on the effect of temperature on sensory rhythms. Proc. Roy. Soc. London B 125: 524, 1938.

(141) SCHOEPFLE, G. M. Synaptic delay and central inhibition in relation to electrotonic potentials. Fed. Proc. 5: 92, 1946.

(142) SMALLWOOD, W. M. AND M. T. HOLMES. The neurofibrillar structure of the giant fibers in Lumbricus terrestris and Eisenia foetida. J. Comp. Neurol. 43: 327, 1927.

(143) STOUGH, H. B. Giant nerve fibers of the earthworm. J. Comp. Neurol. 40: 409, 1926.

(144) TAYLOR, G. W. The optical properties of the earthworm giant fiber sheath as related to fiber size. J. Cell. Comp. Physiol. 15: 363, 1940.

(145) TAYLOR, I. R. AND F. CRESCITELLI. The electrical changes in response to illumination of the dark- and light-adapted eye of Dissosteira carolina. Physiol. Zool. 17: 193, 1944.

(146) THERMAN, P. O. The action potentials of the squid eye. Am. J. Physiol. 130: 239, 1940.

(147) TOBIAS, J. M. AND J. J. KOLLROS. Loci of action of DDT in the cockroach (Periplaneta americana). Biol. Bull. 91: 247, 1946.

(148) WELSH, J. H. The caudal photoreceptor and responses of the crayfish to light. J. Cell. Comp. Physiol. 4: 379, 1934.

(149) WELSH, J. H. AND W. SCHALLEK. Arthropod nervous systems: a review of their structure and function. Physiol. Rev. 26: 447, 1946.

(150) WEVER, E. G. A study of hearing in the sulfur-winged grasshopper (Arphia sulphurea). J. Comp. Psychol. 20: 17, 1935.

(151) WEVER, E. G. AND C. W. BRAY. A new method for the study of hearing in insects. J. Cell. Comp. Physiol. 4: 79, 1933.

(152) WIERSMA, C A. G. Function of the giant fibers of the central nervous system of the crayfish. Proc. Soc. Exper. Biol. Med. 38: 661, 1938.

(153) WIERSMA, C. A. G. The efferent innervation of muscle. Biol. Symp. 3: 259, 1941.

(154) WIERSMA, C. A. G. Giant nerve fiber system of the crayfish. A contribution to comparative physiology of the synapse. J. Neurophysiol. 10: 23, 1947.

(155) WIGGLESWORTH, V. B. The principles of insect physiology. London, Methuen, 1939.

(156) WIGGLESWORTH, V. B. Organs of equilibrium in flying insects. Nature 157: 655, 1946.

(157) WILSKA, A. AND H. K. HARTLINE. The origin of "off responses" in the optic pathway. Am. J. Physiol. 133: 491P, 1941.

(158) WOOLLARD, H. H. AND J. A. HARPMAN. Discontinuity in the nervous system of coelenterates. J. Anat. 73: 559, 1939.

(159) WULFF, V. J. AND T. L. JAHN. Intensity—EMF relationships of the electroretinogram of beetles possessing a visual diurnal rhythm. J. Cell. Comp. Physiol. 22: 89, 1943.

(160) WULFF, V. J. AND T. L. JAHN. The electroretinogram of Cynomya. Anat. Rec. 96: 10, 1946.

(161) YOUNG, J. Z. Synaptic transmission in the absence of nerve cell bodies. J. Physiol. 93: 4P, 1938.

(162) YOUNG, J. Z. Fused neurons and synaptic contacts in the giant nerve fibres of Cephalopods. Phil. Trans. Roy. Soc. London B 229: 465, 1939.